International Journal of American Linguistics

EDITOR
C. F. VOEGELIN
Indiana University

ASSOCIATE EDITORS
PAUL RIVET
Paris

HANS VOGT
Oslo

WILLIAM THALBITZER†
Copenhagen

FANG KUEI LI
University of Washington

REVIEW EDITOR
HARRY HOIJER
University of California; Los Angeles

Volume 24 Number 2

Franz Boas Centennial Volume

April 1958

Published at the Waverly Press, Inc.
Mt. Royal and Guilford Aves.,
Baltimore 2, Md.
by Indiana University
under the auspices of
Linguistic Society of America
American Anthropological Association
with the cooperation of
Joint Committee on American Native Languages

THE INTERNATIONAL JOURNAL OF AMERICAN LINGUISTICS was founded in 1917 by Franz Boas and designed for the publication of data referring to American Indian languages. After forty years of exclusive devotion to those languages, the scope of the Journal is somewhat enlarged: to include papers on languages outside the Americas where such papers illustrate points of theory, method, and technique which have some bearing on American Indian linguistic research. The Editor will invite such contributions and will welcome correspondence from contributors who wish to publish their work in IJAL.

Primary emphasis will, however, continue to be given to data-oriented papers on American Indian languages which may be submitted, without prior correspondence, to the Editor—C. F. Voegelin, Indiana University, Bloomington, Indiana. Books and papers for review, as well as reviews themselves, should be sent to the Review Editor—Harry Hoijer, University of California, Los Angeles 24, California.

JOINT COMMITTEE ON AMERICAN NATIVE LANGUAGES

Harry Hoijer, University of California, Los Angeles (chairman; and representative of the American Council of Learned Societies)

Floyd G. Lounsbury, Yale University (representative of the American Anthropological Association)

Norman A. McQuown, University of Chicago (representative of the Linguistic Society of America)

C. F. Voegelin, Indiana University (editor of the International Journal of American Linguistics, ex officio)

The Joint Committee on American Native Languages serves as the advisory board of the International Journal of American Linguistics and is a body for planning, coordinating, and encouraging research and publications relating to the languages native to the Western Hemisphere.

International Journal of American Linguistics is published quarterly at the Waverly Press, Inc., Mt. Royal and Guilford Avenues, Baltimore 2, Md., by Indiana University. Subscriptions ($5.00 a year) will be received by the Waverly Press, Inc., Mt. Royal and Guilford Avenues, Baltimore 2, Md., or by the Business Manager, Mrs. Elsie Dosch, Indiana University, Bloomington, Ind. Supplements may be ordered separately.

Entered as second class matter May 11, 1945 at the post office at Baltimore, Maryland under the Act of March 3, 1879

International Journal of American Linguistics

VOLUME XXIV April, 1958 NUMBER 2

ALTERNATIVE PHONEMICIZING IN JIVARO

GLEN D. TURNER
INDIANA UNIVERSITY

0. Introduction
1. [e]
2. [z]
3. [ž]
4. (ŋg)[1]

0. The phonemic inventory of Jivaro includes 14 consonants (p, t, k, c, č, s, š, h, m, n, ŋ, w, r, y) and 4 vowels (i, i, a, u) plus stress and nasalization combinable with each vowel.[2] The problem presented here in no way affects the phonemic inventory. It is rather a problem of alternative grouping of a phone with either of two already established phonemes. In each case the phonetic distributions of the phone and of the phonemes involved allow us to analyze the phone as an allophone of either phoneme.

In dealing with each of these alternative groupings we first give in detail (in the subsections marked **.1**) the purely phonetic distributional data pertinent to the phone and phonemes in question. It is on the basis of this data that the phone is seen to be truly alignable with either of the two phonemes in question. In subsection **.2** we examine further evidence for grouping the phone first with one phoneme, then the evidence for grouping it with the other phoneme. In subsections **.3**, **.4**, and **.5** we discuss the alternative possibilities for phonemicization.

1. In a phonetic transcription of Jivaro, [e] occurs commonly. The phonemic problem is whether to group this phone as an allophone of /i/ or as an allophone of /a/. It can easily be shown that the distribution of [e] is mutually exclusive with phonetic [a]. [e] is also in complementary distribution with phonetic [i], but the complementarity is less obvious.

1.1 [e] occurs immediately preceding or following /i/ or /y/ and following, but not preceding, the palatalized allophone of the stops, the nasals, h, w. This statement virtually exhausts the distribution of [e].[3]

[a] also occurs preceding or following /i/ or /y/ but only in the following environments: between /i/ or /y/ and the nonpalatalized allophones [p], [m], [w], and between [p], [m], or word juncture [+] and /i/, /y/. The sequences involved are [iap, yap, iam, yam, iaw, yaw; pai, pay, mai, may, +ai, +ay]. This gives the distribution of [a] when contiguous to /i/ or /y/.

In order to demonstrate that where [e] occurs [a] does not and that where [a] occurs [e] does not, it is necessary to indicate what sequences have not been found. [e] has not been found in any of the sequences given above in the distribution of [a]. [a] has not been found following a palatalized allophone

[1] We use (ŋg) in parentheses to refer collectively to the problematical phonetic sequences [ŋg, ŋgʸ, ŋk, ŋkʸ]. We use (ŋ) to refer to the possible phonemic entity /ŋ/ interpreted to include as allophones the four sequences indicated by (ŋg).

[2] The Jivaro language is composed of four main dialects: Ecuadorian Jivaro, Achuara, Huambisa, and Aguaruna. The Aguaruna are the most divergent, possibly to the point of being a separate language. The data for this paper are from Ecuadorian Jivaro, gathered during 1955–57 under the auspices of the Summer Institute of Linguistics.

[3] In word final position the phonetic sequence [ei] may reduce to [e] with no apparent conditioning factor other than possibly speed of utterance. This does not change the picture as presented in the text.

of any consonant, nor in any sequence with /i/ or /y/ other than those enumerated above. The distribution of [a] is otherwise unlimited.

Since [e] is shown to have one distribution and [a] another, the two are in complementary distribution and therefore groupable as allophones of a single phoneme, phonemic /a/.

There is, however, evidence which alternatively suggests that [e] be grouped with [i] as allophones of phonemic /i/. Though [i] is found neither contiguous to /i/ nor following any of the palatalized allophones, it does occur preceding (and only preceding) /y/. The sequence [iy] occurs when one of the few verb stems which end in [i] is followed by the past tense third person subject marker -yi, e.g. wiyi *he went*. In order to show contrast between [e] and [i] one would look for the sequence [eyi] preceded by a consonant. But the sequence [Ceyi] does not occur. It should also be noted that the sequences [piy] and [miy] have not been found.

Just as phonetic [e] and phonetic [a] are shown above to be groupable as allophones of phonemic /a/, so also phonetic [e] and phonetic [i] are mutually exclusive and could be grouped as allophones of phonemic /i/.

1.2. In section **1.1** we have presented the minimum information necessary to establish each of two alternative phonemicizations involving the phone [e]. In section **1.2** we add further evidence bearing on the problem of [e].

1.2.1. There are two further bits of evidence for considering [e] an allophone of /i/. First, when analyzed this way all possible clusters of two vowels occur. /ai/ would be found with the bilabials as indicated above; [ei] and [ie] would be phonemically /ii/ and /ii/ respectively. Grouping [e] with /a/ would do away with the two clusters /ii/ and /ii/. This argument is weak because the morphophonemics of the language makes those two sequences unexpectable. /i/ plus /i/ becomes /ii/, e.g. núiti *it is that* (from nu *that* plus -iti *it is*) but núkiiti *it is just that* (from nu *that* plus -ki *just, only* plus -iti *it is* morphophonemically become -iti following /i/).[4]

A second bit of evidence for [e] as an allophone of /i/ is that [e] would thus make a more symmetrical arrangement of the phonemes and their allophonic variations in terms of oppositions of high vs. low and front vs. back.

Chart A shows the approximate placement of the four contrasting phones, as follows:

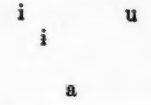
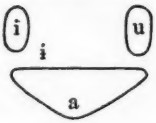


Chart B shows the allophonic range of the vowel phonemes if [e] is considered an allophone of /i/, as follows:

Chart C shows the allophonic range when [e] is grouped with [a] to form /a/, as follows:[5]



1.2.2. Evidence for considering [e] an allophone of /a/ is less tenuous than the evidence given above for [e] as an allophone of /i/.

[4] Inasmuch as there are no true prefixes in the language and none of the suffixes begin with /i/, it is impossible to set up a sequence of /i/ plus /i/.

[5] Simply from an internal study of the distribution of phonemes it would seem that proto-Jivaro may have evidenced only three vowels /*i, *a, *u/ and that in certain distributions /i/ split off from *i. This is uncertain at present and cannot be adduced as evidence, though it would point toward making Chart C more plausible from a historical point of view.

The morphophonemics of Jivaro would be considerably simpler if [e] were considered an allophone of /a/. Morpheme final [a] plus morpheme initial [i] result in [ei]; [i] plus [a] yields [ie]—unless, of course, [a] also follows [p] or [m] (in the case of [ai]) or precedes [p], [m], or [w] (in the case of [ia]). E.g. yąwą́ *dog* plus 3rd person possessive -į becomes yąwę́į *his dog*; cacárį *his mother-in-law* plus -aš *perhaps* becomes cacáriešˇ *his mother-in-law perhaps?* This phenomenon is sufficiently common that it would make the morphophonemics quite clumsy if [e] were grouped with /i/ instead of /a/.

1.3. Among the solutions discussed we choose to make [e] an allophone of /a/. The choice is arrived at by first establishing that on the basis of phonetics alone either solution is conceivable. Once that is established, the added weight of evidence from morphophonemics allows a choice between solutions.

2. [z] is much less common in occurrence in Jivaro than [e]. Though the problem of [z] is similar to that of [e], it is different in one respect. Though [z] is mutually exclusive with [s], every occurrence of [z] may vary freely with [ts]. Thus, [z] is analyzable as an allophone either of /s/ (on the basis of complementary distribution) or of /c/ (on the basis of non-contrastive free variation).

2.1. [z] occurs only following /n/; [s] never occurs there. This shows simple, complementary distribution. However /c/ affects the picture. It is most fortis preceding a voiceless vowel or consonant and in this position it does not vary with [z]. Following /n/, /c/ is lenis (frequently varying with [z]) if also followed by a voiced vowel or /r/. There is never a contrast between [z] and [ts]. Thus, [z] may be grouped equally well either with /s/ or with /c/.

2.2.1. About the only thing to be said for treating [z] as an allophone of /s/ is that it is phonetically similar and mutually exclusive with [s]. This in itself would be ample evidence for grouping them as allophones of a single phoneme were it not for the overwhelming evidence for grouping [z] with [ts] as a single phoneme /c/.

2.2.2. Although [ts] and [z] are somewhat dissimilar phonetically, the free variation between them is one which fits into the allophonic patterns of the language. When following nasal consonants and preceding voiced vowels or consonants, the stops and /č/ have lenis allophones freely varying between voiced and voiceless. It is quite plausible phonetically to explain the variation from [ts] to [z] by saying that the [t] component of the phone [ts] when following /n/ (and only /n/) becomes sufficiently lenis to be assimilated to the homorganic nasal /n/, leaving only the final segment [s] to be voiced. This is particularly expectable since [dz] does not occur in the language. Thus, with [z] as an allophone of /c/ the stops and affricates are symmetrical in their allophonic distribution following /n/.

If one assumes that anything which may be dealt with once and for all in the subphonemic description of a language (without raising problems in the morphophonemics) should be considered sub-phonemic, then one would consider the free variation between [ts] and [z] to be variation between allophones rather than between phonemes. Once described, the free variation need not again be mentioned.

It would seem from the morphophonemics that the fact that [s] and [z] are in mutually exclusive environments is simply an 'accident' of the language. /n/ plus /s/ becomes /nc/; one simply would not expect to find phonemic /ns/.

2.3. [z] is considered an allophone of /c/ on the basis of phonetic symmetry, morphophonemic simplicity, and morphophonemic improbability of any other solution, as shown in the three preceding paragraphs.

3. [ž] occurs even less often than [z]. The problems are quite similar in that [ž] is mutually exclusive with [č] and wherever [ž] occurs it may vary freely with [š]. Here

the similarity ends, however. [ž] and [z] occupy quite different spots in the sound system as will be seen below and the problem of [ž] is not quite so easily resolved as the previous ones.

3.1. [ž] occurs only preceding /t/; [č] never occurs there. Again we have simple, complementary distribution. /š/ is not parallel to /c/ in the previous problem in that in all utterances only one allophone of /š/, namely [š], is heard, without distinction as to fortis or lenis. [š] may occur preceding /t/ but only in non-contrastive free variation with [ž]. Thus, [ž] may be grouped with either /š/ or /č/.

3.2.1. Further criteria for considering [ž] an allophone of /š/ are that they are phonetically more similar than [ž] and [č]; furthermore, on the basis of the assumption above, the free variation between [ž] and [š] should be relegated to the sub-phonemic (allophonic) description since they never occur in contrastive distribution.

3.2.2. The evidence for grouping [ž] with /č/ is peculiar—as languages go—because the only occasion in the language in which [žt] occurs is when the negative allomorph -č- would be expected to occur preceding /t/. From the standpoint of the morphophonemics of the negative morpheme it would simplify the description if [č] and [ž] were grouped as allophones of a single phoneme.

The shift of [č] to [š ~ ž] could be explained phonetically as dissimilation from the /t/ which follows, by loss of the stop component of the [tš] and even occasional voicing of the remaining component [š] (as opposed to the voicelessness of the /t/). This is unparallelled anywhere else in the language however.

3.3. There are actually three alternative solutions for the problem of [ž]. First, one could consider [ž] as an allophone of /š/, leaving to the morphophonemics the task of describing the predictable change of form of the negative morpheme from /-č-/ to /-š-/. Or second, one could consider [ž] an allophone of /č/, leaving to the morphophonemics the describing of free variation between /č/, phonetically [ž], and /š/, phonetically [š]. Or third, one could consider both [ž] and [š] allophones of /č/ before /t/, assuming a sort of intersection of phonemes which would in this case cause no ambiguities and would be equally reversible from phonetic to phonemic transcription and vice versa.

3.4. To deal with [ž] we assume that intersection of phonemes should be avoided if possible even at the cost of more complex morphophonemics. Furthermore, to avoid intersection, the apparently intersecting phones should be checked for fine phonetic distinctions not previously noted. The inference in the second part is that there may be an undetected difference in the [š] allophone of /š/ and the [š] which occurs in free variation with [ž], in which case the third solution for [ž] would be the most acceptable. On the other hand, until such difference is heard (and until now there does not seem to be any difference) the system should be structured so that intersection is avoided inasmuch as there are other alternative solutions presented in the language.

We assumed previously that where free variation is involved, it should be analyzed as sub-phonemic if possible. It is simpler and more consistent to describe in the grammar the phonemically defined alternant of the negative morpheme than to describe freely varying alternation of phonemes and the resultant morpheme alternation. This eliminates solution two in favor of solution one.

4. The problem of (ŋg) differs considerably from the preceding ones. The question here is whether [ŋg, ŋg^{y}, ŋk, and ŋk^{y}] should each be considered a phonetically complex allophone of /ŋ/ (hereafter called the allophone solution), or whether they should be interpreted as a phonemic cluster of /ŋ/ plus various allophones of /k/ (hereafter called the cluster solution). There is nothing

whatsoever in the sound system that would keep one from interpreting (ŋg) either way.

4.1. Five segments in Jivaro, the four above and [ŋ], are all mutually exclusive in distribution. [ŋ] occurs preceding all consonants other than /r/ and in word final position; [ŋg] occurs preceding /r/. In the sequence V_1NV$_2$, N is phonetically [ŋg] when V_1 is any vowel other than /i/ and V_2 any voiced vowel; [ŋg^y] when V_1 is /i/ and V_2 is voiced; [ŋk] when V_1 is not /i/ and V_2 voiceless; [ŋk^y] when V_1 is /i/ and V_2 voiceless. This distribution of itself would be sufficient to group all five phones into a single phoneme. But we must also consider the possibility of interpreting the phonetic sequences as separate phones ([ŋ] and [g]; [ŋ] and [g^y]; etc.) rather than complex unit phones. In this case the separate phones would be allophones of /ŋ/ and /k/ respectively. The cluster solution would leave only one allophone of /ŋ/, namely [ŋ]. However /k/ appears with four allophones [k, g, k^y, g^y] whether or not the cluster solution is adopted. The distribution of the allophones of /k/ is parallel to that of the allophones of /p/ and /t/ and to understand the position of /k/ in the sound system it is necessary to describe the main allophonic distributions of the stops.

Perhaps the simplest method of describing the distributions of the stop allophones is to set up a series of environments, assigning to each an order of priority such that an environment of higher priority precludes the occurrence of an allophone normally expected in an environment of lower priority. Thus, in first priority, when preceding another consonant, only the non-palatalized allophones [p, t, k; b, d, g] occur. Second, preceding voicelessness (voiceless consonants, vowels, or pause) only the voiceless allophones [p, t, k; p^y, t^y, k^y] occur. Third, when preceded by /i/ the voiceless palatalized allophones [p^y, t^y, k^y] occur. When preceded by a nasal consonant the voiced allophones [b, d, g] occur. When preceded by /i/ plus nasal consonant in that order the allophones are both voiced and palatalized [b^y, d^y, g^y]. Thus it is clear that the cluster solution of (ŋg) fits the established patterns of the language; following /ŋ/ the allophones of /k/ would be expected to occur in exactly the distributions in which they actually do occur.

4.2.1. We consider first the evidence for interpreting (ŋg) as a cluster, phonemically /ŋk/. The evidence is good but based almost exclusively on patterning or symmetry.

4.2.1.1. There is no canonical pattern in the language which would cause one to make a preliminary analysis of (ŋg) as unit phones. [ts] and [tš] occur in word initial position and also in word medial position both preceded and followed by other consonants. In both these positions it is necessary to consider them as unit phones in order to preserve the otherwise stable limits of consonant clustering in these positions. This is not the case with (ŋg). They do not occur in an otherwise single consonant position.

4.2.1.2. /ŋk/ may be viewed from two standpoints. It is a sequence of two consonants which demonstrate certain allophonic characteristics when in that particular sequence. It is also a consonant cluster which fits into the overall patterns of nasal-stop consonant clusters. As a sequence and as a cluster /ŋk/ fits into the established sound patterns and fills in gaps which would otherwise be left in the patterns. This evidence is expanded in the next two sections.

4.2.1.3. As first member of the consonant cluster /ŋk/, /ŋ/ shows no peculiarities. Like /m/ or /n/ when occurring as first member of a consonant cluster ending in /k/, the unmodified allophone [ŋ] occurs regardless of what precedes it or what follows the /k/. By contrast with clusters ending in /k/, clusters of nasal plus /r/ automatically insert a stop homorganic to and immediately following the nasal.

As second member of the consonant cluster /ŋk/, /k/ is actualized as [k, k^y, g, g^y] depending on what precedes the /ŋ/ and

what follows the /k/ as noted above. To repeat the distribution briefly as it applies to the sequence /ŋk/, /k/ is voiceless before voiceless vowels, otherwise voiced; it is also palatalized following /i/, i.e. in the sequence /iŋk/. When compared with the allophonic distribution of the other stops, /k/ thus shows perfect symmetry. Grouping (ŋg) as allophones of (ŋ) would not affect the number of allophones of /k/ since the voiced allophones would be evident following /m/ or /n/. To consider (ŋg) a cluster of two consonants simply fills out the distributional picture of /k/ allophones by including /ŋ/ along with the other nasals as being associated with its lenis allophones, i.e. /ŋ/ affects /k/ in the same way /m/ and /n/ do. Thus also, a single distributional statement could be given for all stops which would describe their allophonic environments, i.e. /ŋ/ affects /k/ in the same way it affects /p/ and /t/.

4.2.1.4. By far the most common clusters in frequency of recurrence are those composed of nasal plus homorganic stop. They may occur in morpheme medial or final position, across morpheme boundaries (i.e. morpheme final nasal plus morpheme initial stop), even preceding a third consonant /r/ (forming some of the few three-consonant clusters). These distributions are characteristic of /ŋk/ as a nasal-stop cluster. It would seem strange indeed to find /mp/ and /nt/ occurring frequently with /ŋk/ never occurring once even though /ŋ/ and /k/ occur independently with normal frequency. The allophone solution would define out of existence the sequence /ŋk/.

4.2.2. The evidence for multiple allophones of (ŋ) is also good. It is not completely incontrovertible, however, and in section **4.3** we re-examine the evidence given immediately below in terms of a larger frame of reference. Here we point out certain distributions and symmetrical patterns as they affect /ŋ/; in the next section we discuss the same distributions and patterns as they affect the sound system as a whole or at least parts of the system other than /ŋ/.

4.2.2.1. The allophone solution of (ŋ) reduces the number of consonant clusters without augmenting the phonemic inventory inasmuch as /ŋ/ is undoubtedly phonemic whether or not it is interpreted as having only one allophone [ŋ] or all five allophones [ŋ, ŋg, ŋg^{y}, ŋk, ŋk^{y}]. If all occurrences of the complex phones mentioned are considered allophones of (ŋ), the sequence /ŋk/ is defined out of existence, making one less consonant cluster.

4.2.2.2. With the allophone solution, (ŋ) demonstrates a palatalized allophone [ŋg^{y}] when in the same environment as the palatalized allophones of /m/ or /n/, viz. when preceded by /i/ and followed by a vowel. This symmetry with other nasals would be lost under the cluster solution.

4.2.2.3. With the allophone solution, (ŋ), like all other consonants, would also occur between vowels. The distribution of (ŋ) would be rather circumscribed under the cluster solution.

4.2.2.4. Voiceless vowels in Jivaro present problems which are still unsolved because occasionally—for no apparent reason—they may be completely lost. On such occasions the phonetic change is from [ŋkV] (in which V represents a voiceless vowel) to [ŋ], e.g. [áišmaŋkU ~ áišmaŋ] *man, male*, both of which may occur in isolation with the same linguistic antecedent, for example in answer to the question "What is it?" Morphophonemically this would be simply loss of final vowel if [ŋk] is treated as a unit allophone of (ŋ). Otherwise it would have to be described as loss of final vowel plus simplification of final consonant cluster by loss of the second consonant.

4.2.2.5. The converse of the above is also true. When a morpheme final [ŋ] is followed by a morpheme initial vowel the result is invariably [ŋg-vowel]. With (ŋg) as allophones of (ŋ) there would be no morphophonemic process to explain.

4.2.2.6. The predominent pattern under which metathesis may occur is morpheme

final -VCVC. Metathesis of the final VC gives the bound form -VCCV-. However the form [kutáŋgur] *my bench* plus -ša *and* becomes [kutaŋgrúša] *and my bench* as though the canonical form [kutáŋgur] ended in -VCVC (-aŋur) and thus had metathesized to -VCCV- (-aŋru-).

4.3. The evidence in section **4.2.1** is of such a nature that it does not admit further discussion. It simply states that /ŋk/ as a cluster fits the patterns of the language. But so does the allophone solution. However, the evidence presented in section **4.2.2** for the allophone solution does admit further examination.

Section **4.2.2.1** states that there would be fewer consonant clusters under the allophone solution. But in the evidence for a cluster solution (**4.2.1.4**) it was already stated that the very consonant cluster defined away would be one of the most expectable from the standpoint of other patterns in the sound system.

Section **4.2.2.2** states that /ŋ/ would lose all allophonic symmetry with other nasals unless it could demonstrate a palatalized allophone. Not only is this true but it could be added that though a cluster interpretation would apocopate the subphonemic picture of (ŋ), a unit interpretation would not materially affect the subphonemic picture of /k/ (see **4.2.1.3**). As opposed to this, however, it could be pointed out that even at best there is little to be said for symmetry among nasal consonants. /m/ is more similar to /p/ than to the other nasals in that both /m/ and /p/ are labialized preceding /i/, also in that their distribution with [a] in contrast to [e] is identical. /n/ is broader in terms of the environments in which it is palatalized (both preceding and following /i/) than either /m/ or /ŋ/. In this point /n/ is identical with /t/. Furthermore, in an allophone solution, though /ŋ/ would be more similar to the other nasals in that it now has a palatalized allophone, it would show greater dissimilarity than ever in having five allophones of which only two [ŋ, ŋg^y] would be parallel to /m/ or /n/ in any way.

Section **4.2.2.3** states that /ŋ/ as a phoneme would, under a cluster solution, be limited in distribution inasmuch as it would not occur between vowels. It could be pointed out that at best (ŋ) when in consonant clusters may occur as first member only. Even if limited to a single allophone, the distribution of /ŋ/ could be stated simply as occurring in syllable final position only. This, of course, would not be at all unique among languages of the world as a distributional characteristic of /ŋ/.

Sections **4.2.2.4**, **.5**, and **.6** may be dealt with together in that the same things could be said about certain /mp/ and /nt/ clusters. Examples of loss of final voiceless vowel and the preceding consonant are áyumpA ~ áyum *rooster*; ánkantA ~ ánkan *free*. Examples of nasal plus vowel becoming nasal-stop-vowel are šíam *chick* plus -aš *perhaps* becomes šíampaš *a chick perhaps?*; nuhín *egg* plus the first person possessed suffix -ur becomes nuhíntur *my egg*. Examples of metathesis are ayúmpur *my rooster* plus -ša *and* becomes ayumprúša *and my rooster*; pántam *banana* plus the objective suffix -n becomes pánman, as in the utterance pánman apúhui *he is planting bananas*.

The problem here is much more complex than it appears on the surface. Actually these examples may rather be internal evidence for a set of prenasalized stops which were at one time phonemic entities. Aguaruna still demonstrates phonemic pre-nasalized stops which, though definitely unpredictable and therefore phonemic, seem to be to the Aguaruna informants indistinguishable from the corresponding nasal consonants.

4.4. Assuming for the moment the possibility of a prenasalized stop, at least in velar position, we could still satisfy the evidence on both sides of the (ŋg) question by interpreting any phonetic cluster (ŋg) which occurs within a morpheme as though it were an allophone of (ŋ) and considering as a sequence /ŋk/ those occurrences of (ŋg)

which are divided by morpheme boundary. Compare, for example, the absolute form [kutáŋ ~ kutáŋkA] /kutáŋa/ *bench* with the same stem plus the deictic suffix -ka, [kutáŋgá] /kutáŋká/ *the bench*.

4.5. If, however, one assumes that the phonetic transcription must be transferable per se (i.e. on the basis of purely phonetic phenomena) into phonemic transcription, and vice versa, by means of general distribution statements alone, then the compromise solution for (ŋg) is blocked. It would be impossible to know simply from the phonetic form whether (ŋg) occurred within a morpheme or across morpheme boundaries. Consider the words [kutáŋgá] *the bench* and [tsukáŋgá] *toucan*, the first of which is two morphemes, the second, one.

This eliminates the compromise solution for (ŋg). It would seem that slightly greater clarity of description of Jivaro could be attained with the cluster solution though this difficulty is not insurmountable. Any choice between the cluster and the allophone solutions of (ŋg) would be based on comparatively insignificant details. To be consistent a choice should be made but in this case the choice is arbitrary.

LINGUISTIC PREHISTORY IN THE GREAT BASIN

SYDNEY M. LAMB
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Sub-families
2. The six Numic languages
3. Divergence in the light of lexicostatistics
4. Historical reconstruction

1. The few observations which I am able to make are based on linguistic evidence. If I state some conclusions which seem to conflict with what might be concluded from archaeological and ethnological evidence alone, an opportunity will be provided for discussion and the possibility of a synthesis which might prove to be fruitful.[1]

My observations stem largely from consideration of degrees of linguistic difference among the Great Basin dialects in connection with their relative geographic positions. Therefore, it will be necessary to consider briefly certain aspects of Utaztekan classification.

When Daniel Brinton set up Utaztekan in his classification of 1891,[2] he divided it into three branches, Shoshonean, Sonoran, and Nahuatlan. This three-way division, stemming from a tradition going back to Buschmann's work of a century ago,[3] became traditional itself. Buschmann had set up a linguistic stock consisting of Shoshonean and Sonoran; he held that their resemblances to Aztec were the result of extensive borrowing. Buschmann apparently arrived at the division into Sonoran and Shoshonean largely on geographic grounds; he lacked sufficiently detailed and accurate data to make possible a rigorous linguistic classification.

Later investigations by Kroeber,[4] Mason,[5] and Whorf,[6] led to the realization that the traditional Sonoran family was not actually a genetic linguistic unity, but a geographically-based collection of independent families belonging to Utaztekan. Likewise, Whorf arrived at comparable conclusions with regard to Shoshonean; and his suggestion that it be discarded as a genetic group is supported by lexicostatistic results obtained by Morris Swadesh.[7] My appraisal of the work which has been done on the problem of Utaztekan breakdown suggests that the stock is best divided into some eight or nine families, not just three.[8] This indicates a complex dialect area shortly after the time of Proto-Utaztekan, rather than a

[1] This paper was presented at the Fourth Great Basin Archaeological Conference held at San Francisco in August 1957.

[2] Daniel C. Brinton, The American Race (New York, 1891).

[3] Joh. Carl Buschmann, Die Sprachen Kizh und Netela von Neu-Californien. (Berlin, 1856); Die Lautveränderung aztekischer Wörter in den sonorischen Sprachen und die sonorische Endung AME. (Berlin, 1857); Die Spuren der aztekischen Sprache im nördlichen Mexico und höheren amerikanischen Norden. Zugleich eine Musterung der Völker und Sprachen des nördlichen Mexico's und der Westseite Nordamerika's von Guadalaxara an bis zum Eismeer. (Berlin, 1859); Grammatik der sonorischen Sprachen. (Berlin, 1864).

[4] A. L. Kroeber, Uto-Aztecan languages of Mexico, Ibero-Americana 8 (1934).

[5] J. Alden Mason, The classification of the Sonoran languages, in Essays in Anthropology (Berkeley, 1936) pp. 183–198.

[6] B. L. Whorf, Review of Kroeber, AA 37.343-5 (1935); The comparative linguistics of Uto-Aztecan, AA 37.600-8 (1935); Appendix to Mason, in Essays in Anthropology (Berkeley, 1936) pp. 197–8.

[7] Algunas fechas glotocronológicas importantes para la prehistoria nahua, Revista Mexicana de Estudios Antropológicos 14.173–92 (1954–55).

[8] The term 'family' is used in this paper to refer to subdivisions or branches of larger groupings, which are called 'stocks'. The details of my classification of Utaztecan will be given in a forthcoming paper.

simple splitting into two or three branches. Thus, Kroeber's four divisions of Shoshonean are probably to be considered four independent families of the Utaztekan stock. Consequently, it is not very meaningful, and it is in fact misleading, to speak of Plateau Shoshonean, Pueblo Shoshonean, Kern River Shoshonean, and Southern California Shoshonean, if there is no such thing as Shoshonean. Perhaps I need not mention the fact that these terms are also rather long and cumbersome. In renaming these families, I have followed a suggestion of Gleason, who has proposed that our system of linguistic classification can be clarified by using the suffix '-ic' for subdivisions, or families, of linguistic stocks, reserving the suffix '-an' (or '-n') for the larger groupings. This device conforms to the established tradition in Indoeuropean, where we have Indoeuropean with the suffix '-an', and the subgroups Slavic, Germanic, Italic, and so forth. Pueblo Shoshonean, then, can be renamed simply Hopic, and similarly we have Tubatulabalic. For the Southern California group, we can choose a well-known representative language, such as Luiseño, giving us the name Luisenic. For Plateau Shoshonean a convenient term is that based on the word for person or Indian in these languages, which has the form nïimmi, nimi, niwï, or something similar, depending on the dialect. Thus, the family could be called Numic. In Kroeber's classification of fifty years ago,[9] he divided Numic into three subgroups, which he called Mono-Paviotso, Shoshone-Comanche, and Ute-Chemehuevi. My investigations have indicated that Kroeber was correct, both in considering Numic a distinct genetic group, and in dividing it into these three branches.

2. Apparently the question of just what type of linguistic situation exists within each of these subfamilies has never been settled. During parts of the summers of 1953, 1954, and 1955, I was enabled by the Survey of California Indian Languages of the University of California, to do field work on Mono-Paviotso dialects. I also collected some data on Shoshoni dialects, Kawaiisu, and Tubatulabal. My findings can be briefly summarized as follows: Among the Northern Paiute dialects of Nevada, northeastern California, and Oregon, one finds very slight differences from dialect to dialect; and even dialects as far apart geographically as those of Pyramid Lake and Warm Springs are surprisingly similar to each other. The Bannocks of Fort Hall, Idaho, speak a dialect which differs only very slightly from the dialects of the Paiutes of Oregon and northern Nevada. Moving farther south, the differences seem to be greater for a given distance, but there is nevertheless marked similarity all the way to Mono Lake. Between Mono Lake and the northern Owens Valley, however, the situation is quite different. Here we find a very clear boundary, marked by a number of important differences. The degree of mutual intelligibility is so slight that it is best to consider the dialects as belonging to two separate languages. The language boundary may be drawn along the watershed divide between the headwaters of the Owens River and Mono Lake Basin.[10]

Throughout Owens Valley, Deep Springs Valley, Fish Lake Valley, and the territory of the so-called Western Mono, there are important dialectic differences, but the dialects can be said to belong to a single language. Thus, the Mono-Paviotso branch consists of two languages, one at the southern extremity of the area, the other occupying a much larger area to the north.

This situation raises the question of how these two languages are to be named. The

[9] A. L. Kroeber, Shoshonean dialects of California, UC-PAAE 4.65–165 (1907).

[10] After I had arrived at this conclusion, Professor Kroeber informed me that C. Hart Merriam had made the same observation many years earlier. It apparently did not appear in print, however, until the publication of his posthumous Studies of California Indians (Berkeley, 1955) pp. 149–74.

'Western Monos' are known to themselves and their white neighbors as Monos. In the Owens Valley we have 'Paiutes', known to ethnographers as the Owens Valley Paiute. For the other language, we have the small group known as Bannock, and, again, people called Paiutes. In the interest of following the obviously desirable practice of having a single name for a single language and vice-versa, we must resign ourselves to acceptance of the principle that the name of a language need not coincide with the name of the people who speak it. This situation comes up in more than one place in Utaztekan, and is well known to us in Indoeuropean, where we have people called Americans speaking the English language, Mexicans speaking Spanish, Brazilians Portuguese, and so forth.

The language of Owens Valley and the Sierras can be called either Mono or Monachi. A. L. Kroeber, in a manuscript currently being prepared for publication, has reviewed the history of these names and their use in considerable detail, and has concluded that Monachi is the more desirable of these two names. Versions of the word Mono were used by various Penutian-speaking groups of California to refer to their Numic neighbors to the east, so that the name was applied to speakers of either language, depending on the location of the people using the name. Monachi, a variant of this term with a Yokuts suffix, was applied by the Yokuts to their eastern neighbors, both the Western Mono and the Owens Valley Paiute. It is thus a less ambiguous term than Mono. For the other language, the term 'Northern Paiute', consisting of two words, would be somewhat clumsy as a name for a language, and it might lead to confusion since it is often used, as an ethnographic term, to include the Paiute of Owens Valley. The name 'Paviotso' has been used in the past, and might very well be used again, as a simple, unambiguous name. The word apparently comes from the term used by Nevada Shoshonis for speakers of this language, and the same term, or a very similar one, namely /papioco?o/ is in use in Owens Valley to refer to these people. Since the speakers of this language have no name of their own to set them apart from speakers of other languages, we can do no better than use the name applied by their neighbors, as we have done in the case of Monachi.

One of the most striking characteristics of Mono-Paviotso, or Monachi-Paviotso, is the great disparity in the size of the areas occupied by the two languages. Furthermore, there seems to be at least as much dialectic variation in the small Monachi area as there is in the much larger Paviotso area.

Turning to the Ute-Chemehuevi branch, we find a situation which is surprisingly similar. There seems to be little doubt that all the dialects of the people known as Ute, Southern Paiute, and Chemehuevi are dialects of one and the same language. This language can be called 'Ute' in preference to the longer and potentially confusing alternative name, 'Southern Paiute.' Kawaiisu is markedly different from these dialects, although it is clearly closely related to them. The situation is described by Kroeber, in the manuscript referred to previously, as follows:

"Within Ute-Chemehuevi, Kawaiisu . . . appears slightly but definitely set apart from all other dialects in the division. That is, Chemehuevi, Southern Paiute, and Ute together form a much larger unit whose variations seem to intergrade continuously even when they are fairly marked, whereas Kawaiisu does not participate to the same degree in continuity with its near Southern Paiute neighbors. . . . Kawaiisu did not separate from Las Vegas Southern Paiute with any such recency as . . . the Chemehuevi. . . . The period of their preponderant separateness might be estimated as two or three times as long as that of the Chemehuevi from the Las Vegas Paiute; say roughly about 500 years, plus or minus."

My own guess is that the time period may well be somewhat more than 500 years. At

any rate, it is probably best to speak of Kawaiisu as a separate language.

Again, we have one language occupying a very small territory, and its partner dwarfing it geographically. One is tempted to suppose that perhaps a similar situation exists with regard to Shoshone-Comanche. Unfortunately, I do not yet have sufficient data to definitely ascertain that this is the situation, but the scanty investigation which I have carried out indicates that it probably is. It is quite clear that the Comanches speak the same language as the Shoshonis of Wyoming, Idaho, Utah, and Northern Nevada. This language can of course be called Shoshoni. In the southwestern corner of the Shoshoni territory we have what may be a separate language; it could be called Panamint or Koso, even though the people involved are known as Shoshonis. Panamint informants in Lone Pine told me that they found 'Nevada Shoshoni' difficult to understand, but that they could make out a little bit. This is similar to what I was told by a Monachi informant with reference to Paviotso; it turned out that she was able only to pick out isolated words here and there upon hearing a recorded Paviotso text from Bridgeport. A rough lexicostatistic comparison of the Little Lake and Lemhi dialects gives a time-depth of a little less than eight centuries. Five centuries is a convenient standard to use as a border-line for the difference between dialects of one language and separate languages; so the indication would be that we are dealing with separate languages here also.

If this picture is correct, we have six languages in the Numic family, two belonging to each subdivision. Three of the languages are in a small cluster near the southwestern part of the Great Basin, while the other three occupy a vast area to the north and east. Remembering that the split into the three subfamilies preceded the separation of each of them into a pair of languages, one must place the linguistic center of gravity of the Numic family somewhere around Death Valley. This conclusion is supported by the fact that the closest linguistic relative of Numic is Tubatulabal, directly across the Sierra Nevada from this spot.

3. Kawaiisu and Ute share certain structural features with Tubatulabal which are not found in Monachi-Paviotso nor in Panamint-Shoshoni. And the latter two groups resemble each other slightly more than either of them resembles Kawaiisu-Ute. Nevertheless, Kawaiisu-Ute is clearly much closer to Monachi-Paviotso and Panamint-Shoshoni than it is to Tubatulabal. This situation, like the earlier separation of Utaztekan, can be accounted for by the assumption of a gradual spread into dialects, rather than a series of sharp cleavages.

A few lexicostatistic figures will aid the determination of chronology, but they must be accompanied by the warning that linguists are by no means agreed on the accuracy of lexicostatistic dating. Some linguists contend that lexicostatistic time depths are fairly accurate; others consider them so inaccurate that they have little or no value. My opinion lies between these extremes: I would say that lexicostatistic figures must be considered very rough approximations, subject to such wide margins of error that they can be used only with very great caution, but that they are nevertheless better than nothing at all. With these reservations, we may consider some figures arrived at by Morris Swadesh in 1954.[11] For Monachi and Ute, 19 centuries; Ute and Tubatulabal, 29 centuries; Monachi and Tubatulabal, 35 centuries. This would place the separation of Numic from Tubatulabalic somewhere in the neighborhood of three thousand years ago, followed in about a millenium by the split up of Numic. A few additional figures might

[11] See fn. 7.

be of interest: for Ute and Luiseño, the figure is 38 centuries; for Tubatulabal and Luiseño, 40 centuries; Tubatulabal and Hopi, 33 centuries; Hopi and Papago, 32 centuries; Hopi and Cahita, 31 centuries; Cahita and Aztec, 27 to 29 centuries, depending on the Aztec dialect used. As I mentioned earlier, the split of Utaztekan apparently took the form of a gradual spread into a complex dialect area. Closely adjoining dialects would naturally be more similar than distantly separated ones; and the same situation will hold for the families later developed out of these dialects. The closely adjoining families of Utaztekan now show time depths in the neighborhood of three thousand years. Perhaps the extreme for Utaztekan would be represented by Monachi and Aztec; the time depth of this separation has been calculated by Swadesh as 47 centuries.

4. From the foregoing observations, we may draw some tentative conclusions. They are, of course, to be taken as constituting only a very rough approximation to the true picture. Around five thousand years ago, or somewhat later, Proto-Utaztekan was slowly beginning to spread out into a number of dialects. It may have been located somewhere in the neighborhood of the boundary between Arizona and Sonora. During the second millenium B.C., Numic was becoming distinct from Tubatulabalic. The present location of the surviving handful of Tubatulabal speakers may well be the same as that occupied by their linguistic ancestors of three thousand years ago. The separation of Numic and Tubatulabalic was gradual, and Numic was already beginning to show some dialectic difference before the separation from Tubatulabal became complete. The dialect which was to become the Kawaiisu-Ute branch must have been under the influence of Tubatulabal for some time, while it remained mutually intelligible with the dialect which later separated into Monachi-Paviotso and Panamint-Shoshone. The original Numic dialects may have occupied an area which could have included the historical territory of Kawaiisu as well as the country just east of that part of the Sierra Nevada now assigned to Tubatulabal. The separation of Kawaiisu-Ute and the rest of Numic may be placed in the neighborhood of two thousand years ago; and this was followed very soon thereafter by the separation of Monachi-Paviotso from Panamint-Shoshone. This gives us three Numic languages occupying only a small part of the Great Basin, until perhaps around one thousand years ago. At about this time, for some reason, there began a great movement northward and eastward, which was to extend the domain of Numic far beyond its earlier limits. Since the three branches of Numic were already distinct, this great spread must have been undertaken independently by each of the three groups. The separation in all of them, however, may have taken place at roughly the same time. This would indicate some kind of common influence present in the Great Basin, drawing these people farther into it.[12]

Finally, within each of the three languages Paviotso, Shoshoni, and Ute, dialectic differences are so slight that one can only with the greatest difficulty imagine that they could have occupied the vast areas in which we find them for more than a very few centuries. Since these languages spread out from a southwestern corner, the most likely conclusion, from linguistic evidence, is that much of the northern and eastern part of the Great Basin was not occupied by speakers of the present Numic languages at the time Columbus discovered America. And, as of around 1000 years ago and earlier, the major part of the Great

[12] Linguistics apparently does not provide the answer as to what kind of influence this might have been. A suggestive possible correlation is the distribution of the bison, shown in Julian Steward, BAE-B 120, p. 37.

Basin is unaccounted for linguistically. This area may have been occupied by speakers of languages which moved elsewhere, or the languages may have become extinct. Some of the languages which became extinct could, of course, have been Utaztekan, but they could just as well, or perhaps more likely, have been languages related to Hokan, Zuni, Keres, Algonkian, or even some stock now totally extinct. The people who inhabited this large unaccounted-for area may have moved out, or they may have remained, and adopted the new languages.

INTERNAL DIVERSITY IN UTO-AZTECAN: I

KENNETH HALE
UNIVERSITY OF NEW MEXICO

0. Introduction
1. Languages used in test
2. Method of cognate determination
3. Separation dates and groupings

0. The purposes of this paper are (1) to group, on the basis of lexicostatistics, a sampling of Uto-Aztecan languages and (2) to present a method for determining whether the items in a given comparison are cognate.

Lexicostatistics, variously called glottochronology or the Swadesh method, is a method whereby the time of separation of related languages can be ascertained in an approximate number of years. A basic list of 215 relatively culture free items was first developed which exhibited, on an average, 81 retentions out of 100 for a period of 1,000 years. By the law of chance, the percentage of retention for two related languages of the same time level was 81% of 81%, or 66%.[1] Later, in 1955, a new list of 100 items was developed which exhibited a higher rate of retention (r = 86%) which, for two related languages of the same time level, raised the retention rate per 1,000 years to 86% of 86%, or 74%.[2] In the present paper, the second list was used. To determine the time of separation (denoted by t) of a pair of languages, the list was translated into the two languages, and the lists were compared to determine the percentage of cognates (denoted by C). The date was then worked out on the basis of the formula: $t = \log C \div \log r^2$.[3]

[1] Morris Swadesh, Lexico-Statistic Dating of Prehistoric Ethnic contact, APSP 96.452–463 (1952).

[2] Morris Swadesh, Towards Greater Accuracy in Lexicostatistic Dating, IJAL 21.121–137 (1955).

[3] Ibid.

1. Seventeen languages of the Uto-Aztecan stock were examined for the present study. The languages were selected purely on the basis of availability; no attempt was made to concentrate on any one area or to pregroup the languages on the basis of work already done in the Uto-Aztecan field.

The following languages were used: Northern Paiute (Np), Southern Paiute (Sp), Ute (U), Shoshone (Sh), Comanche (Cm), Tübatulabal (Tu), Cahuilla (Ca), Hopi (Ho), Papago (Pa), Pima (Pi), Tepecano (Tp), Mayo (Ma), Tarahumara (Ta), Cora (Co), Póchutla (Po), Mecayapan (Me), Zacapoxtla (Za).[4]

[4] I wish to express my gratitude to the authors of the various sources which I used to obtain data for the test lists. I also want to thank Stanley Newman for his criticism and help.

The data are taken from the following sources: Np—from a tape of Swadesh's Penutian Survey List in the Archives of the Languages of the World, Indiana University; Sp—Edward Sapir, The Southern Paiute Language, Proceedings of the American Academy of Sciences 65.1–730 (1930); U—J. P. Harrington, The Phonetic System of the Ute Language, University of Colorado Studies 8.199–222 (1911): Sh—D. B. Shimkin, Shoshone I: Linguistic Sketch and Text, IJAL 15.175–188 (1949); Shoshone is also taken from tape in my possession; Cm—Joseph B. Casagrande, Comanche Linguistic Acculturation, IJAL 20.140–151, 217–237, 21.8–25 (1954–1955); Tu—Charles F. Voegelin, Tübatulabal Grammar, UCPAAE 34.55–190 (1935); Ca—A. L. Kroeber, Shoshonean Dialects, UPCAAE 8.235–269 (1909); Ho-Pa-Pi—from data of my own; Tp—J. A. Mason, Tepecano, A Piman Language of Mexico, Annals of the N. Y Academy of Science 25.309–416 (1916); Ma—Howard Collard and J. Dedrick, Cahita, Instituto Lingüístico de Verano, 1955 (Swadesh list, unpublished); Ta—K. S. Hilton, Tarahumara, Instituto Lingüístico de Verano, 1955 (Swadesh list, unpublished); Co—K. Preuss, Wörterbuch Deutsch-Cora, IJAL 8.81–102; Po—F. Boas, El Dialecto de Póchutla, Oaxaca, IJAL 1.9–43 (1917);

The test lists for the various languages were obtained mainly from published word lists, grammars and texts. For inclusion of a language in the test, a lower limit was arbitrarily set at 70 vocabulary items in the Swadesh 100 word list. Due to the rather low limit, certain pairs of languages yielded dates of separation which were inconsistent with other pairs. But it was hoped that the number of languages tested would act as a corrective measure so that diagramming of the splits would be possible regardless of certain aberrant dates. This proved to be the case.

Changes in the original transcriptions were introduced only to make the data more uniform and easier to print. No attempt was made to phonemicize data which was not already phonemic.

2. Items in a given comparison were marked plus on the basis of attestable phonetic correspondences. A tentative reconstruction of proto-Uto-Aztecan was made using only the data available in the test lists; the relevant parts of compared items were marked plus only if they exhibited reflexes of a common proto-form. No improvement of reconstructions by inclusion of data outside the lists was made.

The value of this method lies in the fact that a controlled method of cognate determination may be applied when available data are limited to those which occur in the Swadesh list. The ideal situation, in which a great deal is known of the comparative linguistics involved, is seldom forthcoming for most of the languages of the world.

By use of this method, it is also possible to isolate borrowings, coincidental similarities, and irrelevancies. For example, Np tohá- *white* and Pa túha *white* are considered too uncertain to be accepted as cognate, since only t and a are attestable correspondences. Two items exhibiting a difference in length are considered cognate if part of one of the items shows correspondence to a part or the whole of the other. Thus, Ca puš *eye* and Za iištootoolo *eye* are considered cognate, since puš and iiš- exhibit a one to one correspondence. This solution is enhanced by the comparison of Me iiš *eye* to the Za form.

2.1. A tentative reconstruction of proto-Uto-Aztecan phonology is presented in this section. Because of the limited range of vocabulary which could be admitted as evidence, and because of the variety of transcriptions in which the evidence occurred, it has been necessary in certain instances to accept unconditioned multiple reflexes of a single proto-phoneme. Two consonants in a given daughter language whose difference is one of a single manner of articulation have been admitted as possible reflexes of a single proto-phoneme. Unconditioned multiple vowel reflexes may differ in tongue height from high to mid or from low to mid, but not from high to low. Vowel reflexes may also differ in tongue position from front to back and in manner from rounded to unrounded. Consonantal correspondences are given higher rank than vowel correspondences. If, in a given comparison, all the consonants correspond and some or none of the vowels do, the items are considered cognate, but if all the vowels correspond and some or none of the consonants do, the items are considered too uncertain to be included as cognates. Thus Pa háahagi *leaf* and Me šiwit *leaf* are considered cognate, since the relevant parts show consonantal correspondence. That is, the morphology of the Pa form shows the initial syllables háaha- of which h corresponds to š in Me which includes a final -t, presumably a suffix, external to the comparison; but Pa -g- corresponds to Me -w-.

Each of the following reconstructions is followed by its reflexes in the daughter languages. Multiple reflexes are followed,

Me—H. W. Law, Nahua de Mecayapan, Instituto Lingüístico de Verano, 1955 (Swadesh list, unpublished); Za—Harold and Mary Key, Vocabulario Mejicano de la Sierra de Zacapoxtla, Puebla, 232 pages, 1953.

where possible, by a brief, parenthetic reference to the conditioning factor. The numbers at the end of each paragraph refer to the reconstructed forms in **2.2.**

Consonants: *p: Np p; Sp p; U p; Sh p; Cm p; Tu p (initially), b (finally); Ca p; Ho p; Pa-Pi w (initially), p (medially); Tp b, w (initially), p (medially and finally); Ma p, b (initially), bb (medially following single v́); Ta b; Co h; Po Ø, Me Ø (initially), p (medially); Za Ø. Examples: 1–5.

*t: Np t; Sp t (initially), r (medially); U (initially), r (medially); Sh d (initially, r (medially); Cm t (initially), r (medially); Tu t (initially), l (medially); Ca t (before a, u), č (before e); Ho t (before a, i), q (before ö); Pa-Pi t (before a, o), č (before i, i, u); Tp t, D; Ma t (initially), tt (medially); Ta r, t; Co t; Po t; Me t; Za t. Examples: 5–18.

*k: Np k (initially), q (medially); Sp q (initially and following consonants), γ (medially); U k; Sh g (before a), k (elsewhere); Cm k; Tu h (initially and medially), g (finally); Ca k (initially and medially), x (finally); Ho q (before a), k (before i); Pa-Pi k; Tp k; Ma k, kk; Ta k; Co k; Po k; Me k (initially), ʔ (finally); Za k. Examples: 10, 20, 22–26.

*c: Ho c; Ma c; Co c; Po c; Me c; Za c Examples: 19, 49.

*q: Np q; Sp q; Tu k; Ho k; Pa-Pi-Tp k; Ma k; Ta k; Za k. Example: 27.

*kʷ: Np kʷ; Sp qw; U kʷ; Sh gʷ, kʷ; Cm kʷ; Ca qw; Ho kʷ; Pa-Pi-Tp b; Ma bʷ; Ta w; Co k (before i), kʷ (elsewhere); Po kʷ; Me kʷ; Za kw. Examples; 12, 19, 28, 29.

*ʔ: Np ʔ; Sp ʔ, Ø; U ʔ, Ø; Sh Ø; Ho ʔ; Pa-Pi ʔ; Tp ʔ, Ø; Ma ʔ; Ta Ø; Co Ø; Po Ø; Me Ø; Za Ø. Examples: 16, 30–32.

*s: Np s; Sp š, s·; U s (initially), š (medially); Sh s (initially), š (medially); Cm s; Tu š; Ca s (initially), š (finally); Pa h; Pi h (initially), Ø (medially); Tp h (initially), Ø (medially); Ma s (initially), ss (medially following single v́); Ta s; Co s; Po s; Me s, š; Za s, š. Examples: 4, 13, 28, 33–36.

*h: Np h; Sp Ø; U ʔ (initially and between identical v), Ø (medially); Sh h; Cm h; Tu Ø, ʔ; Ca h (initially), Ø (medially); Ho h (initially), Ø (medially); Pa-Pi-Tp h (initially before a), ʔ (initially before i), Ø (medially); Ma h; Ta h; Co h (before a): Ø (before i); Po Ø; Me Ø; Za Ø. Examples: 14, 30, 31, 37, 38.

*m: Np m; Sp m (initially), ŋw (medially); U m (initially), ŋw (medially); Sh m; Cm m; Tu m; Ca mʔ; Ho m; Pa-Pi-Tp m; Ma m; Co m; Po m; Me-Za m (initially), n (finally). Examples: 9, 15, 18, 23, 24, 35, 39, 40, 43.

*n: Np n; Sp n (initially), m (before bilabial stops), n· (medially between vowels); U n; Sh m (before bilabial stops), n (elsewhere); Cm n; Tu n; Pa-Pi n (before a, o), ŋ (before i, i, u); Tp n; Ma nn (medially following single v́), n (elsewhere); Ta n; Co n; Po n; Me n; Za n. Examples 7, 21, 41–44, 49, 50.

*ŋ: Sp ŋ; U ŋ; Sh n; Tu ŋ; Ca mʔ; Ho ŋ (initially), m (medially); Pa-Pi n; Tp N; Na n (initially), nn (medially); Ta n; Co n; Me n; Za n. Examples: 8, 45.

*w: Np w; Sp w; U w; Sh w; Cm w; Tu w; Ca w; Ho l (before a, ö), w (before i, i); Pa-Pi g; Tp g (initially), G (finally); Ma w; Ta Ø (before o), w (elsewhere); Co w; Me w; Za w. Examples: 26, 32, 34, 44, 47, 48.

*y: Ho y; Pa-Pi d (before a), ǯ (before u); Tp d; Ma y; Ta Ø; Po y; Me y; Za y. Examples: 22, 25.

In the following list of vowel reconstructions, length and stress (both are probably phonemic in all the languages examined) are not marked, since it has not been considered necessary for the purposes of this paper. In citing the reflexes in the daughter languages, length and stress have been omitted.

*i: Np i; Sp i; U i; Sh i; Cm i; Tu i; Ca i; Ho i; Pa-Pi-Tp i; Ma e (following k or h), i (elsewhere); Ta u (following k), i (elsewhere); Co i; Po i; Me i; Za i. Examples: 1, 9, 12, 19, 23, 25, 26, 28, 48.

*e: Np i, u; Sp i; U i, u; Sh i, i, e, ë; Cm

i, u; Tu i; Ca e, u; Ho i; Pa-Pi-Tp ɨ; Ma e; Ta i; Co e; Po e, u; Me e; Za e. Examples: 4, 10, 16, 17, 27, 33, 40, 41, 43, 44, 47.

*ɛ: Ca a; Pa-Pi i; Tp u; Ma i; Co a; Po e; Za e. Example: 50.

*a: Np a; Sp a; U a; Sh a; Cm a; Tu a; Ca a; Ho a; Pa-Pi i (finally following stops), a (elsewhere); Tp a; Me e (following y), a (elsewhere); Ta a; Co a; Po a (finally), e (elsewhere); Me a; Za e (following y), a (elsewhere). Examples: 2, 6, 9, 10, 12, 13, 15, 16, 18, 20–22, 24, 26, 28, 29, 31, 32, 34, 35, 37, 39, 45.

*u: Np i, o; Sp u; U i, u; Sh u; Cm u; Tu u; Ca u; Ho o; Pa-Pi-Tp u; Ma u; Ta u; Co i; Po e, u; Me i; Za i. Examples: 1, 14, 23, 25, 36.

*o: Np ɔ; Sp ɔ; U o, u, i; Sh ɔ; Cm o, u; Tu o; Ca i; Ho ö; Pa u (either contiguous to a, or separated from a by ʔ, h), o (elsewhere); Pi-Tp o; Ma o; Ta o, u; Co i, u; Po u; Me o; Za o. Examples: 3, 5, 13, 30, 31, 39, 42.

*ɔ: Np a; Sp a; U a; Sh a; Cm a; Tu o; Ca a; Ho a, ö; Pa-Pi-Tp o; Ma o; Ta o; Co u; Me a; Za a. Examples: 8, 46.

Clusters: *mk: Sp ŋʔw; U ʔw; Sh ŋgw; Ho q; Po k; Me g; Za k. Example: 6.

*nk: Np q; Sp ŋq; U ŋk; Sh ŋk; Cm k; Tu ŋh; Ca qʔ; Ho q; Pa-Pi k; Ma kk; Ta k; Co k; Po k; Me k; Za k. Example: 21.

*nt: Np Ø; Sp m; U Ø; Sh m; Cm Ø; Tu nt; Ma tt; Ta t; Co t; Po t; Me t; Za t. Example: 11.

*ns: Np Ø; Sp ʔ; U ʔ; Sh ʔ; Cm Ø; Tu nʒ; Ca š; Ho s; Pa h; Pi Ø; Tp ʔ; Ma s; Ta s; Co s; Po š; Me š; Za š. Example: 1.

2.2. There follow, in this section, 50 reconstructed items. Each of the reconstructions serves merely as an index of correspondences held in common by a number of cognate items. They are not reconstructed forms in the usual comparative sense, since total sound control was, in most cases, impossible due to limitations inherent both in the data and in the method. Each reconstructed item in the following list is followed by its counterpart in the daughter languages in which it occurs.

1. *punsi *eye*: Np pɨi, Sp puʔi, U piʔi, Sh púʔI, Cm pui, Tu punʒi-, Ca puš Ho póosi, Pa wúhi, Pi wúi, Tp buʔi, Ma puúsi, Ta busí, Co hisi, Po -štotolú, Me -iiš, Za iištootoolo.

2. *pa *water*: Np pá·, Sp pa·, U paA, Sh pa·, Cm pá·, Tu pa·l, Ca pal, Ho páahi, Ma baáʔa, Ta baʔwí, Co he, Po at, Me aat, Za aat.

3. *po *path*: Np pó·, Sp pɔ·, U piɬ, Cm puʔE, Tu poh-, Ca pit, Ho pôhi, Pa wóogi, Pi woogi, Tp voi, Ma boóʔo, Ta buwé, Co huye, Po otkán, Me ihti, Za ohti.

4. *sepe *cold*: Sp šipi, U si, Tu ši·b-, Pa-Pi hɨ́ipi, Tp hipit, Ma sébbe, Co se, Me sekti, Za seseek.

5. *top *fish*: Pa-Pi wátopi, Tp vatop, Me topo.

6. *tamka *man*: Sp taŋʔwa, U taʔwa, Sh dangwəp, Cm tenapimaʔ, Tu ta·twal, Ho táaqa, Cora ta·ta, Po teket, Ma tagat, Za taakat.

7. *ten *mouth*: Np tupá, Sp tímpa, U tipa, Sh dimbI, Cm tu·pE, Pa-Pi čɨ́ni, Ma teéni, Ta riní, Co teni, Po ten, Me -ten, Za teen.

8. *toŋ *knee, leg*: Sp taŋa-, U tá·ŋa-, Sh danëp, Tu toŋo·-, Ca tamʔi, Ho támö, Pa-Pi tóoni, Ma tónno, Ta ronó, Co tunú, Me -tankʷa, Za taancwaay.

9. *mati *know*: Tu ma·g, Pa-Pi máači, Tp má·tik, Po metí, Me -mati, Za -mati.

10. *kate *sit*: Np katí, Sp qari-, U kariɬ, Sh gári, Cm karɬ, Tu hal-, Ho qáti, Co ka·teI.

11. *tent *stone*: Np tipí·, Sp timpi-, U tupi-, Sh dimbI, Cm ti·pI, Tu tínt, Ma tétta, Ta rité, Co tete, Po tot, Me tet, Za tet.

12. *kʷi- *earth*: Tp bi·D, Ma bʷíya, Ta wiʔyé.

13. *tosa *white*: Sp toš·a-, U tuša, Cm tosa-, Ho qöcá-, Pa túha, Pi tóa, Tp toa, Ma tósali, Ta rosákami.

14. *tuh *black*: Np toho-, Sp tu·-, U tukʷa-, Sh túhu-, Cm tu·-, Ca tul-, Pa-Pi čúki, Tp tuk, Za tiiltik.

15. *tama *tooth*: Np ta·má, Sp taŋwa,

U taŋwa, Sh damA, Cm tama-, Ca tam?a, Ho táma, Pa-Pi táatami, Tp tamo, Ma támmi, Ta ramé, Me -tan, Za taan.

16. *?ate *louse*: Ho ?ati, Ma ?étte, Po atómt, Me atimit.

17. *te- *small*: Np tɨ́·ci?, Sh dé·či, Cm tie-, Ma teébe, Po túči.

18. *tam, -t- *we*: Np támi, Sp tám·i-, U tami, Sh damë, Cm tami, Ca čemem, Ho ?itámi, Pa-Pi ?áačimi, Tp atiV, Ma ?ítapo, Ta tamuhé, Co itam, Po tuen, Me tehmen.

19. *kʷici *smoke*: Np kʷí·dahA, Sp qwi·-, U kʷipɨ, Sh kʷipu, Cm kʷipɨ, Ho kʷíi?ciŋʷi, Ma bʷicía, Co kicí.

20. *ka *not*: Np kái, Sp qa, U ka, Sh ke, Cm ke·, Tu haš, Ca ki?i, Ho qa, Ma kaa, Ta ke-, Co ka.

21. *nanka *ear*, *hear*: Np naqá, Sp naŋqa, U naŋka, Sh náŋk, Cm na·kI, Tu nanha, Ca naq?a, Ho náqvi, Pa-Pi náaki, Ma nákka, Ta naká, Co naka, Po nekéz, Me nakas, Za nakas.

22. *yaka *nose*: Ho yáqa, Pa-Pi dáaki, Ma yékka, Ta aká, Po yekét, Me -ya?, Za yekacol.

23. *muki *die*: Tu mu·g, Ca mukiš, Ho móoki, Pa-Pi múuki, Tp mui, Ma muúke, Ta mukú, Po mok, Me miki, Za -miki.

24. *maka *give*: Sp maqa, Cm máka, Tu maha, Ca max, Ho máya, Pa-Pi máaki, Tp ma·, Ma miſka, Po meká, Me -maga, Za maka.

25. *yuki *rain*: Ho yóoki, Pa-Pi ǯúuki, Tp du·ki, Ma yúke, Ta ukuí, Po yekt.

26. *waki *dry*: Tu wa·g, Ho láaki, Pa-Pi gáki, Ma wakía, Ta wakicéami, Co waci, Me waki.

27. *qe *bite*: Np qɨ́i, Sp qi?i, Tu ki·?, Ho kɨ́iki, Pa-Pi-Tp kɨ́i, Ma ké?eye, Ta ikí, Za kikeeh.

28. *kʷasi *tail*: Np kʷasí, Sp qwas·i-, U kʷasí, Sh gʷáši, Cm kwasI, Ca qwasqa, Pa báhi, Pi bái, Tp bai, Ma bʷassía, Ta wasí, Co kʷasi.

29. *kʷa *eat*: Ma -bʷa, Ta ko?wá, Co kʷa, Po kʷa, Me kʷaa, Za takwaa.

30. *?oho *bone*: Np –?hó, Sp ɔɔ-, U o?ovI, Ho ?ôôqa, Pa-Pi ?óo?oo, Tp ?ó?o, Ma ?ótta, Ta ocí, Co ika, Po ot, Me oomit, Za oomit.

31. *?oha *yellow*: Np ?ahá-, Sp ɔa-, U ?oa-, Sh zha, Cm ?oha, Pa ?úami, Pi ?oámi, Tp oam.

32. *?awa *horn*: Npwa·, Sp ?a·pi, U ?a·, Cm ?a·-, Tu a·-, Ho ?aala, Pa-Pi ?áa?agi, Tp ?a?aG, Ma ?áa ?am, Ta a?wá, Co awa.

33. *se *one*: Np sɨ·nʷi, Sp ro·-, U su-, Sh sim, Cm si·, Ca supli, Ho sɨ́ika, Pa-Pi hímakoo, Tp hímaD, Ma seénu, Co sai, Po se·, Me se, Za see.

34. *sawa *leaf*: Pa-Pi háahagi, Ma sáuwa, Ta sawá, Me šiwit, Za aašiwit.

35. *masa *feather*: Ho mása, Ta masá.

36. *su *star*: Tu šu·l, Ca suwet, Ho sóohi, Me sitalin, Za siitalin.

37. *ha- *what*, *who*: Np haká, Sp aŋa, an·ia, Sh hága, Cm hakai, Tu agí, Ca haxi, Ho háki, Pa hásičuu, Pi hásiǯuu, Tp hatoV, ha·stu, Ma hábe, Co haní, haiki, Po ak, Me a?kon, Za aakoni.

38. *hi *drink*: Np hívi, Sp ivi-, U ?ivi-, Cm hi·vI, Tu ?ɨ·-, Ho híiko, Pa-Pi ?ii, Tp ?i, Ma hé?oye, Ta bahí, Co i.

39. *ma, mo *hand*: Np ma·i, Sp mɔ?ɔ, U mi?i, Sh mo, Cm mo?o, Tu ma·-, Ca ma?, Ho má, Ma mámma, Co moáka, Po mai, Me maa, Za maay.

40. *me *kill*: Tu mi?ig-, Ca meka, Tp mɨ́?ak, Ma mé?a, Ta mi?yá.

41. *ne *I*: Np nɨ·, Sp ni?, U ni?, Sh nit, Cm ni?u?, Tu nik, Ca ne, Ho ni?, Pa-Pi ?áaŋi?i, Ma ?ínapo, Ta nihé, Co ni, Po nen, Me neh, Za neh.

42. *no *egg*: Cm no?, Ho nôhi, Pa-Pi nóhnaa,

43. *nema *liver*: Np nɨ·mi, Sp niŋwi, U nupI, Sh ni?wi, Cm nim, Tu nim-, Ca nem?a, Ho nɨ́ima, Pa-Pi nɨ́mi.

44. *wene *stand*: Np winɨ́?o, Sp win·ɨ́, U wini-, Sh wënë, Tu wíni-, Ho wini-, Ta wirí.

45. *ŋa *root*: Ho ŋáhi, Ma náawa, Ta nawá, Me nelwayo, Za naalwal.

46. *wɔ *two*: Np wa·hái?i, Sp wa·-, U waiyɨinI, Sh wadë, Cm wáha-, Tu wo-, Ca wi, Ho lôôyi, Pa-Pi góoki, Tp gó·k, Ma woóyi, Ta okuá, Co wapo.

TABLE 1. Uto-Aztecan Dates

	Za	Me	Po	Co	Ta	Ma	Tp	Pi	Pa	Ho	Ca	Tu	Cm	Sh	U	Sp
Np	4213	4595	3992	3217	4099	3298	3702	3779	3206	2954	3206	3039	1046	1427	1328	1748
Sp	4106	4733	3992	3298	4107	3298	3504	3779	3298	2740	2802	2649	1092	1198	618	
U	4267	4214	3992	3038	3992	3038	4588	3481	3389	2879	3389	2649	954	1481		
Sh	5023	5152	4595	3481	4733	3617	4108	4107	4107	2725	3046	2229	424			
Cm	4733	4733	4099	3389	4259	3481	3519	3481	3298	2504	2641	2298				
Tu	4221	3809	3809	3046	3504	3206	3809	3617	3481	2878	2229					
Ca	3617	4466	3992	3671	3481	3672	3481	3298	3122	2878						
Ho	3885	4107	3779	3779	3992	3389	4214	3617	3779							
Pa	3916	3809	3916	3389	3298	3046	427	198								
Pi	4099	3702	4351	3581	3206	2878	695									
Tp	3481	3481	3617	3389	2878	2878										
Ma	3617	3908	2877	2878	3031											
Ta	4466	4595	3779	3779												
Co	4344	3885	3885													
Po	1252	1214														
Me	1145															

47. *we *big, long*: Ho wíiko, wíipa, Pa-Pi gí?iši, Tp gí, Me weyi, weyat· Za weeyi, Za weeyak.

48. *wi *grease*: Tu wípt, Ca wi, Ho wíhi, Pa-Pi gíigi, Ta wi?í.

49. *con *head, hair*: Ma coónim, Po con, Me conkal, Za -coon.

50. *nɛn *tongue*: Ca naŋ, Pa-Pi níjni, Tp nunu, Ma nínni, Co nanu, Po nenepíl, Za nenepiil.

3.1. The dates of separation (in terms of years ago) for the 17 languages examined are contained in Table I.

Many of the inconsistent dates can be explained as being due to faulty data. The Sh test list, for example, is too brief to render consistent dates. Such inconsistencies can be readily found and accounted for. A much more disturbing problem occurs, however, when a group of languages exhibits dates which are consistently out of line. A basic assumption dictates that two closely related languages or groups of languages must be equidistant in time from a language or group of languages which is more distant in time from each of the two closely related languages than the latter are from each other. In this study, however, it has been found that Np, Sp, U, Sh, Cm, Tu, Ca and Ho (Shoshonean) exhibit a consistently lower cognate per cent, yielding an average date of 4100, when compared to Za, Me, and Po (Nahuatlan) than do Pa, Pi, Tp, Ma, and Co (Sonoran) which yield an average date of 3800 for that comparison. But Sonoran exhibits a higher cognate per cent, yielding an average date of 3400, when compared to Shoshonean than when compared to Nahuatlan. The dates, 4100 and 3800, which, theoretically, should be the same, exhibit a discrepancy of 300 years. This discrepancy is probably not great enough to cause any concern, but, at least in the author's mind, the tenet that the rate of morpheme decay is constant and the same for different languages must be taken more liberally than he had hitherto supposed.

Another disturbing problem is presented by the Ta dates. Ta consistently exhibits a low cognate per cent for all of its comparisons except those with Pa-Pi-Tp (Pima-Tepecano) and Ma (Cáhita). The low cognate percents yield dates which are high enough to suggest that Ta constitutes a separate sub-stock equal to the Northern and Southern sub-stocks (see diagram below). The high cognate percents, however, yield dates which suggest that Ta is a Sonoran language. If Ta is Sonoran, then

the premise that different languages lose vocabulary items at the same rate is false. If Ta constitutes a separate sub-stock, then a great deal of borrowing must have occurred between Ta and Sonoran at a date before the phonological changes peculiar to Ta took place.

These problems make diagramming difficult and, in certain cases, wholly arbitrary.

3.2. The diagram in this section gives the groupings of the 17 Uto-Aztecan languages and the approximate dates of separation as indicated by the study. The major split occurred approximately 4,000 years ago dividing the Uto-Aztecan stock into two major sub-stocks, the Northern and the Southern. The Northern sub-stock is further divided into the Shoshonean and the Sonoran. The Shoshonean is divided into the Plateau Shoshonean (Np, Sh, Cm, Sp, U), the Kern River (Tu), the Pueblo (Ho), and the Southern Californian (Ca).

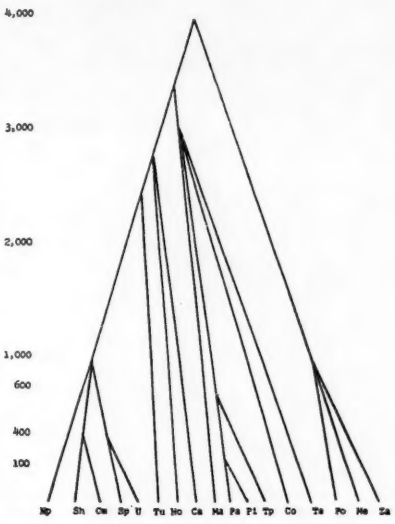


DIAGRAM OF UTO-AZTECAN GROUPINGS

MORPHOLOGICAL STRUCTURE OF ISTHMUS NAHUAT

HOWARD W. LAW

SUMMER INSTITUTE OF LINGUISTICS

0. Introduction
1. Brief description of phonemes
2. Morphophonemics
3. Word classes
3.1. Verbs
3.2. Nouns
3.3. Pronouns
3.4. Particles
4. Derivation
4.1. Verb derivation
4.2. Noun derivation
4.3. Particle derivation
5. Morpheme index

0. The Isthmus dialect belongs to the traditional Nahuat division of the Nahuatl language by virtue of the presence of the sound represented by the letter t instead of the tl of some other dialects. Within the Isthmus dialect there is some minor variation in sounds and vocabulary used, but together the villages in the southern part of the state of Veracruz where Nahuat is spoken compose a speech community contrasting on the dialect level with other Nahua or Nahuatl speaking communities in Mexico. The village of Mecayapan, Ver., is one of a dozen or so where this dialect is spoken, and is the municipio of the area where the forms described in this paper are heard. Other Nahuat speaking municipios in the southern part of the state of Veracruz are: San Andrés Tuxtla, Cosoleacaque, Oteapan, Jáltipan.

The 1940[1] census figures place the number of Nahuat speaking residents in the state of Veracruz at 76,765. No figures are available for the speakers of the Isthmus dialect alone. There is a high degree of bilingualism in many sections, particularly near the coast; and Spanish is becoming understood and spoken by more Aztecs at a rapidly increasing rate especially among the men and boys.

The following description is made on the basis of field trips to the area over a period of nine years beginning in November 1944. It is not intended as a complete grammar. The description of the phonemes is given in condensed form. A fuller description appeared in Mexico Antiguo, Vol. VIII, 1955, pp. 267f. No description of the syntactic structure is included, though such a paper is in process of preparation.

The writer wishes to acknowledge gratefully his appreciation especially to William Wonderly and Robert Longacre as well as the others of his colleagues of the Summer Institute of Linguistics whose encouragement and helpful suggestions have made possible the completion of this paper.

1. Nahuat exhibits the following phonemes: consonants /p, t, k, b, d, g, m, n, c, č, k^w, s, š, h, l, w, y/; vowels /i, e, a, o/; suprasegmental phonemes /ˑ/ (vowel augment) and /ˈ/ (stress).

/p, t, k/ are simple voiceless stops; /b, d, g/ are simple voiced stops. Before a different stop in a word medial cluster, the voiceless stops are aspirated. /t/ and /k/ occur unreleased in word final position. /b/ and /d/ are statistically infrequent and contrast with /p/ and /t/ in intervocalic position only in native Nahuat words. /g/ is somewhat more frequent and contrasts in both intervocalic and word initial positions.

Of the nasal series /m/ has two allophones; bilabial [m] occurs in word initial and intervocalic positions; velar [ŋ] occurs in word final position and before velar stops

[1] The 1950 census records, though available, do not give the figures for the Nahuat speaking population of the state of Veracruz.

and /w/. The contrast between /m/ and /n/ is neutralized in clusters when the nasal occurs before stops or spirants, in which case the nasal is always homorganic with the following consonant.

/c, č, kʷ/ are complex voiceless stops. /c/ is an alveolar affricate, /č/ is an alveopalatal affricate, and /kʷ/ is a labialized velar stop. /kʷ/ and /k/ are neutralized in word final position and as first member of a consonant cluster in which case only /k/ appears.

/s, š, h/ are spirants of alveolar, alveopalatal and back positions respectively. The point of articulation of the /h/ varies according to the tongue position of the vowels contiguous to it. The voiceless component present in the /h/ phoneme extends to the first part of the following voiced continuant /n, l, w/.

/l/ is a voiced alveolar lateral phoneme.

/w, y/ are voiced semivowels. /w/ is phonetically both bilabial and velar. In word medial position after /i/, /w/ may have a velar voiced fricative approach.

/ʔ/ is a voiceless glottal stop and shows contrast in a very restricted number of environments. In word final position it appears only in rapid speech replacing word final /t/ or /k/ of slow speech.

/i/ has an allophone [ɪ] in close syllables before /k/ and, in certain forms, before /h/.

/a/ has an allophone [æ] occurring as a free variant in a few words in some idiolects.

Stress is unmarked when it occurs on the penult of the word unit; it is marked by an acute accent when it occurs on a different syllable of the word.

2. This section of Morphophonemics is limited to alternations pertinent to more than one class in the language. The discussion of changes pertinent to only one class (word class or distribution class within a word class) will be found at the beginning of the discussion of that class.

Clusters of nasals (identical and non-identical) and clusters of identical voiceless stops may reduce in rapid speech to a single consonant of the quality and quantity of the second member: kan 'locative' + más *more* (Spanish loan) > kamás *the most* (*as to place*) e.g. kamás takoˑmik *the darkest place* (takoˑmik *dark*); točam *our home* + 22.1 -meh 'noun plural' > točameh *our homes* (the significance of the morpheme numbers is explained in **3.1.3.1**); 122.3 am- 'second person exclusive[2] plural subject' + motalih *sit down* > amotalih *you* (*pl.*) *are sitting down*; nik- (from 113.3 ni 'first person subject' + 133.4 ki- 'third person object'; see **3.1.3.2.2**) + -kowa *buy* > nikowa *I buy it.*

3. The four word classes described below may be distinguished structurally by the following patterns of formation:

Verbs (see **3.1** and **4.1** below) occur with a prefix of decade 10, prefixes of class 100 and suffixes of class 200, i.e. affixes indicating mode, subject, object, tense, aspect and number.

Nouns (see **3.2** and **4.2** below) occur with morphemes of decades 20–50 and prefixes of class 100, i.e. affixes indicating number, diminutive, vocative, locative and possession.

Pronouns (see **3.3** below) occur with suffixes of decade 60, i.e. affixes indicating plurality and emphasis.

Particles (see **3.4** and **4.3** below) occur uninflected.

3.1. Verbs are treated in regard to (1) Person morphemes, (2) Modal morphemes, (3) Tense morphemes, (4) Root classes, (5) Aspect morphemes and (6) Plural morphemes.

As indicated by the following minimum-

[2] The terms exclusive and inclusive as used in this paper have the following meanings: exclusive—excluding the person speaking; inclusive—including the person speaking, since the exclusion is from the point of view of the hearer, not the speaker as traditionally defined.

maximum formulas there are three prefix orders and two suffix orders to the verb. Of these, one prefix and one suffix are obligatory in declarative constructions. In the following description of verb constructions all the prefixes are discussed first, then the tense morphemes, which are suffixes and obligatory to the verb stem in declarative constructions and which are also determiners for the verb root classes. After the tense morphemes the verb root classes are described. The rest of the verb suffixes, all non-obligatory, are discussed last in the section.

3.1.1. The following summary of verb expressions is given here as a preview of the Nahuat verb before the more detailed distributional discussion which follows. There are two types of minimum-maximum verb expressions—one for each type of verb construction: permissive and imperative, and declarative types.

The minimum verb expression for the permissive and imperative constructions (minVxP-I) is composed of a verb root (V) plus either morpheme 12 ma- 'permissive' and one of the following: 113.3, 121.3, 122.3 or 133.3, or morpheme 126.3 ši- 'imperative subject indicator.' This is expressed in formula as follows: minVxP-I = + (+12 + 100.3)/(+126.3) + V, in which / indicates 'or'.

The minimum verb expression for the declarative verb constructions (minVxD) is composed of a verb root plus a person subject prefix (Class 100 morpheme) plus either a tense suffix (Class 210 morpheme) or an aspect suffix (Class 220–30 morpheme). This is expressed in formula as follows: minVxD = +100.3 + V + (210)/(220–30), in which / indicates 'or'.

The maximum verb expression for the permissive and imperative constructions (maxVxP-I) may be expanded by one or two proclitics plus morpheme 242 -kan 'imperative plural' plus up to five enclitics. This is expressed in formula as follows: maxVxP-I = ±procl. + (+12 + 100.3)/(+126.3) + Vs ± 242 ± encl., in which procl. indicates proclitic, encl. indicates enclitic, and / indicates 'or'.

The maximum verb expression for the declarative constructions (maxVxD) may be expanded by one or two proclitics, a modal prefix (morpheme 11), an object prefix (Class 100 morpheme), a plural suffix (morpheme 241) and up to five enclitics. This is expressed in formula as follows: maxVxD = ±procl. ± 11 + 100.3 ± 100.4 + Vs + (210)/(220–30) ± 241 ± encl., in which procl. indicates proclitic, encl. indicates enclitic, and / indicates 'or'.

3.1.2. At the juncture of a prefix and a stem, clusters of two vowels whether identical or not, except when the first is long (e.g. ne·-+-asi > ne·asi), become a single short vowel of the quality and quantity of the second vowel:[3] 113.3. ni- *I* + -ia *go* > nia *I go*, 121.3 ti- *you (sg.)* + -asi? *arrived* > tasi? *you (sg.) arrived*, 121.3 ti- *you (sg.)* + -ono? *are* > tono? *You (sg.) are*, 121.3 ti- *you (sg.)* + -a·taneki *are thirsty* > ta·taneki *you (sg.) are thirsty*.

3.1.3. Class 100 morphemes are the person markers of the verbs, nouns and pronouns. The patent phonological similarity of the forms: neh *I* (pronoun), no- *my* (noun prefix), ni- *I* (verb subject prefix), and ne·- *me* (verb object prefix) plus the common semantic component of first person allows these forms to be united into a single morpheme numbered 113. Similarly the same factors in the forms: teh, to-, ti-, te·-; ameh-, amo-, am-; and yeh, (i- ~ iy-), ZERO, ki- with the common semantic components of second person plural, second person inclusive,[4] and third person respectively allows each of these groups of forms to be united into single morphemes numbered respectively 121, 122, and 133.

Each, therefore, of the following mor-

[3] There is one exception in normal speech: mo- + -ita > moita *it is seen, he sees himself*, etc. Further exceptions characterize slow or precise speech.

[4] See footnote 2.

phemes of Class 100 has several allomorphs which may occur as a pronoun root, in noun possessor position, in verb subject position and in verb object position: 113 'first person', 121 'second person inclusive', 122 'second person plural', 133 'third person'. Other morphemes of Class 100 have only one allomorph each: 120 'second person singular' occurring as 120.2 mo- 'noun possessor prefix', 125 'second person indeterminate number' occurring as 125.4 mic- 'verb object prefix', and 126 'imperative mode second person' occurring as 126.3 ši- 'verb subject prefix'. Since the majority of the person morphemes have, among other variants, two allomorphs occurring with verb roots, and since the only allomorphs of two others (125 and 126) occur with verb roots, the complete distribution of Class 100 morphemes 'person markers' is given here. The class meaning of the morphemes of Class 100 is 'person indicator' (i.e. in contrast to morphemes of Class 210 and Class 220–30 with class meanings of 'tense' and 'aspect-mode' respectively).

3.1.3.1. The symbolism of the semantic components is here discussed. In the Class 100 morphemes there is a series of three semantic components constituting the primary category of person; these are represented by decade numbers: 10 first person; 20 second person; 30 third person.

Another series of three semantic components appears in combination with the preceding series and constitutes a further category of person which is present or absent in the .1, .3 and .4 allomorphs of Class 100 morphemes dependent upon the surrounding morphemic environment (present only when a plural morpheme occurs with the stem). In the case of .2 allomorphs (see below) the immediate or non-immediate context determines whether this further category of person is present. These further components are represented by unit numbers: 01 first person; 02 second person; 03 third person.

The unit number 01 'first person', when in combination with a decade number, marks a further semantic component of first person present in the morpheme. Examples: tehamen (121.1[5] plus a plural morpheme) *we/us*, i.e. *you and I/me, you and we/us*, točam (121.2 plus a noun root) *our home*, i.e. *your and my home, your and our home*, tikawih (121.3 plus verb root plus a plural morpheme) *we remain*, i.e. *you and I remain, you and we remain*, tikasih (121.3 plus 133.4 plus verb root plus a plural morpheme) *we find it, we find them*, i.e. *you and I/we find it/them*.

The unit number 02 'second person', when in combination with a decade number, marks a further semantic component of second person potentially present in the morpheme. Examples: amehwa·n (122.1 plus a plural morpheme) *you* (*pl.*) i.e. *you and you*, amočam (122.2 plus a noun root) *your* (*pl.*) *home*, i.e. *your and your home*, amkawih (122.3 plus verb root plus a plural morpheme) *you* (*pl.*) *remain*, i.e. *you and you remain*, amkasih (122.3 plus 133.4 plus verb root plus a plural morpheme) *you* (*pl.*) *find it, you* (*pl.*) *find them*, i.e. *you and you find it/them*.

The unit number 03 'third person', when in combination with a decade number, marks a further semantic component of third person potentially present in the morpheme. Examples: nehamen (113.1 plus a plural morpheme) *we/us*, i.e. *I and he, she, it; me and him, her, it;* nočam (113.2 plus a noun root) *my home, our home, i.e. my and his, her, its home;* yehamen (133.1 plus a plural morpheme) *they/them*; ičam (133.2 plus a noun root) *his, her, its home, their home*, kawih (133.3 plus verb root plus a plural morpheme) *they remain*, kasih (133.3 plus 133.4 plus verb root plus a plural morpheme) *they find it/them*.

The combination of a decade number and a unit number with a century number (i.e. Class 100) yields a formula which indicates

[5] For an explanation of the decimal numbers used see below in the same section.

the semantic components of any given person morpheme. The following seven combinations actually occur:

113 *I/me* (and also *they/them*) = *I/me* or *we/us exclusive.*

120 *your sg.* = *your sg.*

121 *I/me and you sg. or pl.* = *we/us inclusive.*

122 *you* (*and also you*), *your pl.* = *you pl., your pl.*

125[6] *you sg.* (and also *you pl.*) = *you sg.* or *you pl.*

126 *you sg.* (and also *you pl.*) (subj.) = *you* (*sg. or pl.*) (subj.).

133 *he/him, she/her, it* (and also *he/him, she/her, it*) = *he/him, she/her, it* or *they/them.*

Most of these semantic combinations are each associated with a set of allomorphs; each set indicates a meaning corresponding to the sum of the semantic components indicated in the combination which thus formulaically summarizes the meaning of the morpheme. The several allomorphs of the above morphemes occur according to the distribution of these morphemes with different root classes and with affix 241. Accordingly each allomorph is numbered according to the following decimal system: .1 pronoun root; .2 noun possessor; .3 verb subject prefix; .4 verb object prefix. (Phonologically determined allomorphs in this paper are symbolized in parenthesis with the symbol ~ between each two allomorphs of the morpheme.) The seven combinations of numbers referred to above (113, 120, 121, 122, 125, 126 and 133) are used along with these decimal numbers to indicate allomorphs of Class 100 morphemes.

3.1.3.2. In this section the morphemes of Class 100, person morphemes, and their allomorphs are listed, discussed, and then described in terms of their distribution. The distribution of the allomorphs depends upon the word class to which each allomorph belongs in that the allomorphs of Class 100 morphemes occur in several different word classes.

3.1.3.2.1. The morphemes of Class 100 and their allomorphs are here listed. Morphemes containing semantic component 03, with their allomorphs, are as follows:

Morpheme 113 'first person'[7] (containing the obligatory component 10 and the optional component 03 and therefore potentially plural): 113.1 neh pronoun root; 113.2 no- noun possessor prefix; 113.3 ni- verb subject prefix; 113.4 ne·- verb object prefix.

Morpheme 133 'third person' (containing the obligatory component 30 and the optional component 03 and therefore potentially plural): 133.1 yeh pronoun root; 133.2 (i- ~ iy-) noun possessor prefix; 133.3 ZERO verb subject prefix; 133.4 ki- verb object prefix.

Morphemes containing semantic component 02 consist of only morpheme 122.

Morpheme 122 'second person plural' (containing obligatory components 20 and 02 and therefore consistently plural): 122.1 ameh- pronoun root; 122.2 amo- noun possessor prefix; 122.3 am- verb subject prefix[8].

Morphemes containing semantic component 01 consist of only morpheme 121.

Morpheme 121 'second person inclusive'[9] (containing the obligatory component 20 and the optional component 01 and therefore potentially plural; allomorphs 121.2 and 121.4 are always plural): 121.1 teh pronoun root; 121.2 to- noun possessor prefix; 121.3 ti- verb subject prefix; 121.4 te·- verb object prefix.

Morphemes containing no unit semantic component consist of morphemes 120, 125, and 126.

Morpheme 120 'second person singular'

[6] The unit number 5 has no semantic significance. This is also true of other numbers above three in the remainder of this section.

[7] Morpheme 113 is inherently inclusive.

[8] Morpheme 125 mic- 'second person object prefix' is treated below.

[9] See footnote 2.

(containing only the obligatory component 20): 120.2 mo- noun possessor prefix.

Morpheme 125 'second person indeterminate number' (containing the obligatory component 20 and an extra-systematic component 05 'indeterminate number'): 125.4 mic- verb object prefix.

Morpheme 126 'second person imperative mode' (containing the obligatory component 20 and an extra-systematic component 06 'imperative mode'): 126.3 ši- verb subject prefix.

Morphemes containing no unit semantic component nor decade semantic component, but belonging to the Class 100, person marker, morphemes, consist of morphemes 145 and 146.

Morpheme 145 ta- distributive verb object prefix 'third person indeterminate number' generalizes the object of the verb or distributes the action of the verb. It is the same order prefix as the .4 allomorphs of the Class 100 morphemes described above, and is mutually exclusive with each of them. It is also mutually exclusive with morpheme 146 described below. It may occur following the .3 allomorphs of the Class 100 morphemes and immediately preceding the verb stem. It also occurs as a verb stem formative in noun derivation (see **4.2**).

Morpheme 146 mo- 'reflexive indicator' (and 'passive voice indicator' with some roots) occurs in all persons and in singular and plural number constructions. It is mutually exclusive with all .4 allomorphs of the Class 100 morphemes and with morpheme 145. It occurs in the same order as morpheme 145.

3.1.3.2.2. Some distributional-semantic considerations regarding the above listed Class 100 morphemes are here given.

Morpheme 122 is different from morphemes 113, 121 and 133 in that morpheme 122 always marks plural whereas morphemes 113, 121 and 133 may mark either singular or plural. Allomorph 122.1 ameh- always occurs with either morpheme 62 or morpheme 63 both indicating plurality with the resultant partial redundancy of both allomorph 122.1 and either morpheme 62 or 63. (Allomorph 122.2 amo- 'second plural' never occurs in redundant distribution with any other morphemic form marking plural.) Morphemes 113, 121 and 133 may also occur with plural morphemes (e.g. the .1 allomorphs of these morphemes occur with either morpheme 62 or 63, -amen ∞ -waˑn and -emen respectively 'noun pluralizers'; the .3 allomorphs of these morphemes occur with morpheme 241 -h ∞ -keh 'verb pluralizers'; the .2 allomorphs of these morphemes never occur with any pluralizer). However, there is no redundancy when such plural indicators occur for the person indicator does not in itself mark plurality, but merely marks a second or added potential category of person.

Allomorph 122.3 am- 'second person plural' and allomorph 121.3 ti- 'second person inclusive' contrast in meaning when they occur in sequence with the plural morpheme 241 -h ∞ -keh. 122.3 am- means *you pl.*, i.e. *you and you*, whereas 121.3 ti- means *we/us*, i.e. *you and I/me*: amkičiˑwah *you pl. are doing it*, tikčiˑwah *we are doing it*, amkipiakeh *you pl. have it*, tikpiakeh *we have it*.

Morpheme 120, morpheme 125 and morpheme 126 are still different from the above four in that morpheme 120 indicates simply the second person singular possessor of the noun and is never pluralized, morpheme 125 indicates the second person verb object with the number not formally indicated in the verb sequence, and morpheme 126 occurs with a special plural indicator, morpheme 242 -kan (see **3.1.7**). In addition morphemes 120, 125 and 126 have only one allomorph each.

Morpheme 120 in its .2 allomorph mocontrasts with the other members of the noun possessor series (allomorphs of morphemes 113, 121, 122 and 133) in the following expressions: nočam *my home, our home*, i.e. *his, her, its and my home*, močam *your*

(*sg.*) *home*, amočam *your* (*pl.*) *home*, točam *our home*, i.e. *your and my home*, ičam *his, her, its, their home*.

Morpheme 125 in its .4 allomorph mic- contrasts with the other members of the verb object series (allomorphs of morphemes 113, 121 and 133) and morpheme 145 ta- and morpheme 146 mo- in the following expressions: ne·maga? *he hit me*, mic- maga? *he hit you*, te·maga? *he hit us*, ki- maga? *he hit him* (*or her or it*), tamaga? *he hit* (*in general*), momaga? *he hit himself*, etc.

Morpheme 126 in its .3 allomorph ši- contrasts with the other members of the verb subject series (allomorphs of morphemes 113, 121, 122 and 133) in the following expressions: šikči·wa *do it*, nikči·wa *I do it*, tikči·wa *you sg. do it*, kiči·wa *he does it*, amkiči·wah *you pl. do it*.

3.1.3.2.3. The distribution of the allomorphs of Class 100 morphemes is as follows:

The sequence of 113.3 ni-, 121.3 ti- or 126.3 ši- plus 133.4 ki- results in loss of the final /i/ of the sequence: ni- + ki- > nik- as in nikmaka *I am giving it to him*, nikpopowa *I am relating it*, etc.; ti- + ki- > tik- as in tikmaka *you are giving it to him*, tikpopowa *you are relating it*, etc.; ši- + ki- > šik- as in šikmaka *give it to him*, šikpopowa *relate it to him*, etc.

All .1 allomorphs occur as free pronoun roots. The .1 allomorphs of morphemes 113, 121 and 133 may occur without a plural morpheme, in which case they mark singular, or in sequence with suffix morpheme 62 or 63 'plural'. (See **3.3.2**). Examples: neh *I/me*, nehamen *we/us*, teh *you sg.*, tehamen *we/us*, yeh *he/him, she/ her, it*, yehamen *they/them*, nehemen *we/us*, yehemen *they/them*.

All the .2 allomorphs are first order prefixes indicating the possessor of the noun. They may be preceded only by proclitics. They are mutually exclusive with each other. The sequence of 113.2 or 133.2 with a noun stem is ambiguous as to the number of the possessor. Examples: nočam *my home, our home*, ičam *his home, their home*, nopilowa·n *my children, our children*, ipilowa·n *his children, their children*.

All .3 allomorphs are second order prefixes indicating the verb subject. They may be immediately preceded only by Class 10 morphemes. A second morpheme of Class 100 in its .4 allomorphic form may appear between the .3 allomorph and the verb stem. All .3 allomorphs are mutually exclusive with each other. Examples: nikawi *I remain*, nikmana *I cook it*, tikawi *you* (*sg.*) *remain*, tikmana *you* (*sg.*) *cook it*, kawi *he remains*, kimana *he cooks it*. The .3 allomorphs of 113, 121 and 133 may occur without a plural morpheme, in which case they mark singular, or in sequence with morpheme 241 'subject plural'. Examples: nikawi *I remain*, niweci *I fall*, nimicita *I see you*, tikawi *you* (*sg.*) *remain*, micita *he sees you*, nikawih *we remain*, tikawih *we remain*, nimicitah *we see you*, niwecih *we fall*, micitah *they see you*.

All .4 allomorphs are first order prefixes indicating the object of the verb. They may be immediately preceded by Class 100 morphemes in their .3 allomorphic form. They may be followed immediately only by the stem. All .4 allomorphs are mutually exclusive with each other and with morphemes 145 and 146. The sequence of 113.4 or 133.4 with a verb stem is ambiguous as to the number of the object. Examples: nikmaka *I am giving it/them to him/them*, etc., kimaga? *he hit him/them* etc., ne·maga? *he hit me/us*, etc.

3.1.4. The following modal indicators of the verbs are morphemes of Class 10. The morphemes of Class 10 are prefixes of the third order from the stem. They may be preceded by only one or more proclitics. They may be followed by the .3 and .4 allomorphs of Class 100 morphemes. They are mutually exclusive with each other.

Morpheme 11 a- 'negative': ania *I'm not going*, anikneki *I don't want it*, akimaga? *he didn't hit him*, ayahkih *they didn't go*.

Morpheme 11 in verb constructions is always atonic as seen in the following: awél *he can't*, ayá *he is not going* (homophonous with ayá *not*), awi? *he is not coming*. The negative particle amo occurs with the imperative mode and precedes the verb construction with which it occurs.

Morpheme 12 ma- 'permissive' indicates action done with the permission of the person referred to, or action done in dependent relation to some preliminary action: manipašalowa *with your permission I visit*, manimictehkawilti *let me lift you up on it*, makiči·wa *let him do it*, kikowa iga makikʷa· *she bought it in order that he might eat it*.

3.1.5. Class 200 morphemes are suffixes to verbs and indicate tense (Class 210), aspect or mode (Class 220–30) and plural (Class 240).

The tense indicators of the verbs, morphemes of Class 210, are suffixes of the first relative order from the stem. They may be followed by the plural morpheme only, morpheme 241 -h ∞ -keh. They are mutually exclusive with each other and with morphemes of Class 220–30. The Class 210 morphemes are here described with their allomorphs:

Morpheme 211 (-ZERO ∞ -a?) 'present'. The present tense indicates action in the present or near present (cf. Spanish present tense usage). The distribution of the allomorphs of morpheme 211 along with those of 212 below are the major determiners for the verb root classes discussed below. Allomorph 211.1 -ZERO occurs with Root Classes I–III; allomorph 211.2 -a? occurs with Root Class IV. Examples: tiwecka *you laugh*, tona *it is hot*, kikowa *he buys it*, niktali *I put it*, kiknelia? *he pities him*.

Morpheme 212 (-? ∞ -h) 'punctiliar past'. The punctiliar past tense indicates action complete at a point in time past. Allomorph 212.1 -? occurs with Root Class I; allomorph 212.2 -h occurs with Root Classes II–IV. Examples: kisa? *he left*, asi? *he arrived*, kisaloh *he stuck it on*, kipiloh *he hung it*.

Morpheme 213 -aya 'incompletive past'. The incompletive past tense indicates action incomplete in the past. It is in contrast to morpheme 212 'punctiliar past' and morpheme 214 'durative past' (see below), with the /i/ final roots of Classes I and III and all roots of Class IV (which are /i/ final also). With the /a/ final roots of Classes I and III and all roots of Class II morpheme 213 -aya indicates the durative past tense also. Examples: kisaya *he was leaving, he was going to leave*, asiaya *he was about to arrive*, panowaya *he was passing by, he was about to pass by*, kikʷa·ya *he was eating it, he was going to eat it*, pahtiaya *it was about to heal*, koniaya *he was going to drink it*.

Morpheme 214 -a? 'durative past'. The durative past tense indicates action going on over a period of time in the past. Morpheme 214 -a? occurs only with Classes I, III and IV. (See above in reference to morpheme 213.) Examples: asia? *he was arriving (and continued to arrive)*, pahtia? *it was healing (and continued to heal)*, konia? *he was drinking it (and continued to drink it)*.

Morpheme 215 -s 'future'. The future tense indicates action in the future. Morpheme 215 occurs with all root classes. Examples: čokas *he will cry*, asis *he will arrive*, kiči·was *he will make it*, patanis *he will fly*, kikʷa·s *he will eat it*, konis *he will drink it*.

In root or stem plus suffix constructions vowels of identical quality (whether long or short) reduce to a single vowel of the same quality as the first, e.g. -maka- *give* + -aya 'incompletive past' > -makaya *was giving*, -či·wa- *do, make* + -aya 'incompletive past' > -či·waya *was doing, was making*, -kʷa·- *eat* + -aya 'incompletive past' > -kʷa·ya *was eating*; -pia- *have* + -a? 'present' > kipia? *he has it*.

3.1.6. The occurrence of the allomorphs of morphemes 211 and 212 are the principal determiners of the four classes of verb roots mentioned above. Class II roots, however, have a third determiner in the occurrence of stem formative morpheme 311

between the root and the first order suffix, thus distinguishing Class II roots from Class III roots. There is a further class, Class V, composed of four miscellaneous roots two of which occur with suppletive stems in some constructions.

Allomorphs 211.1 -ZERO and 212.1 -ʔ determine Class I roots[10] wecka *laugh*, keca *arise, stop*, asi *arrive*, tona *be hot*, and others. Allomorphs 211.1 -ZERO and 212.2 -h, preceded by morpheme 311 -wa ∞ -wi 'stem formative', determine Class II roots: picko *bruise*, ka *leave*, and others. Allomorphs 211.1 -ZERO and 212.2 -h alone determine Class III roots: ma·ma· *carry a load*, patani *fly*, and others. Allomorphs 211.2 -aʔ and 212.2 -h determine Class IV roots: oni *drink*, tani *earn*, and others. Class V is composed of only the following four roots, two of which occur with suppletive stems in some constructions: wi *come* (walah *came*, wiceh *will come*), ia *go* (yah *went*). The root pia *have* occurs with allomorphs 211.2 -aʔ and 212.1 -ʔ with resulting homophonous forms in the present and punctiliar past tenses: kipiaʔ *he has it, he had it*. The root ono *be* occurs with allomorph 211.2 -aʔ in the present and with a unique allomorph -ya in the incompletive past tense instead of allomorph 213.2 -aya, but it is defective in that it has no punctiliar past tense.

3.1.7. The aspect indicators of the verbs, morphemes of Class 220–30 (which extend, contrary to the usual arrangement, from 221–239), are suffixes of the first relative order from the stem. They may be followed by only the plural morpheme, 241, -h ∞ -keh. They are mutually exclusive with each other and with the Class 210 morphemes discussed above.

In the Class 220–30 morphemes there is a series of two semantic components that mark structural categories of aspect; these are represented by decade numbers: 20 complete aspect; 30 incomplete aspect.

Another series of six semantic components appears in combination with the preceding series and also marks a second, but different, structural category of aspect or mode type; these components are represented by unit numbers: 01 return of actor to former state or location; 02 no return of actor to former state or location; 03 action at location of object of the verb; 04 action away from the location of the subject of the verb; 05 frustrative mode; 06 conditional mode.

The combination of a number from the series of two components (decade numbers) plus a number from the series of six components (unit numbers) with a number of a morpheme of Class 220–30 results in an indication of the semantic components of any given morpheme of Class 220–30. The following eight combinations actually occur:

221 -ka complete aspect and return of actor to former state or location.

222 -ki complete aspect and no return of actor to former state or location.

223 -ko completed action at the location of the object of the verb.

224 -to completed action away from the location of the subject of the verb.

233 -ki incompleted action at the location of the object of the verb.

234 -ti incompleted action away from the location of the subject of the verb.

[10] Root Class I may be subdivided by the following indicated determiners:

Subclass A occurs with no change in the stem except optional apocopation in the past completive tense of the final stem phoneme or two phonemes preceding the plural morpheme: -keca- *arise, stop*, -asi- *arrive*, and others.

Subclass B occurs with obligatory change in the past completive tense stem in one of the following categories: 1—Voicing of medial consonant and in certain stems optional apocopation of the final stem syllable: -čoka- *cry*, -maka- *give*, and others. 2—With apocopation of the stem final syllable and velarization of the resulting final consonant: -mayana- *be hungry*, -čihčimi- *touch*, and others. 3—With only final syllable apocopation: -wahliga- *bring*, -witeki- *beat*, and others.

Subclass C is composed of roots which normally only occur in third person constructions: -tona- *be hot*, -tako·mi- *be dark*, and others.

235 -skia? incompleted action frustrated.
236 -skiaya incompleted action conditioned.[11]

The distribution of the above Class 220–30 morphemes is as follows:

Morpheme 221 -ka occurs with all verb roots: yahka *he went* (*and has returned*), kočka *he slept* (*and is awake*).

Morpheme 222 -ki occurs only with ia *go* in the present corpus: yahki *he went* (*and hasn't returned yet*).

Morpheme 223 -ko occurs with all verb roots:[12] nimicitako *I came to see you*, micihliko *he came to tell you*, kitahtaniko *he came to ask for it.*

Morpheme 224 -to occurs with all verb roots: nikawato *I went to leave it*, tikowato *you went to buy it*, tak^{w}a·to *he went to eat* (*a meal*).

Morpheme 233 -ki occurs with all verb roots: ne·itaki *he is coming to see me*, šine·-pašalowiki *come to visit me.*

Morpheme 234 -ti occurs with all verb roots:[13] nikmagati *I am going to hit him*, yati *he is going to go*, nikowati *I am going to buy it.*

Morpheme 235 -skia? occurs with all verb roots: tama·ma·skia? *he came to load, but didn't*, niwiciskia? *I was going to come, but didn't.*

Morpheme 236 -skiaya occurs with all verb roots: nimayanaskiaya *I might be hungry*, tikwahligaskiaya *you might bring it.*

In the description presented here the morphemes which determine verb root classes (Class 210) are grouped together into one major class and the morphemes of Class 220–30 into another.

An alternative and equally valid analysis would group classes 210 and 220–30 into a single class and sort out meaning components of Class 210 as is done for Class 220–30. It would then be necessary to set up two subclasses: members of the class containing allomorphs by which verb roots are classified, other members of the class. These two subclasses would correspond exactly to classes 210 and 220–30. The criterion of classification used in such an analysis is, of course, the mutually exclusive distribution of the morphemes involved.

3.1.8. The plural indicators of the verbs are morphemes of Class 240.

In all verb stems the number is implicitly singular[14] unless otherwise marked by a plural morpheme. Allomorph 122.3 is completely redundant as to the number component in that 122.3 indicates both plural and second person and also occurs obligatorily with morpheme 241 -h ∞ -keh 'plural' in all occurrences. There is no redundancy

[11] Several further analyses are possible for morphemes 223, 224, 233, 234, 235 and 236. The forms -ko, -to, -ki, -ti present obvious recurring partials which even have some meaning correspondences; the forms -skia? and -skiaya have certain recurring partials and corresponding meanings, and in addition are abnormally long for the usual Nahuat morphemes. However, it is difficult to support any division smaller than those described in this paper. All potential breakdowns result in some subanalysis comparable to isolating the sl of slip, slosh, slush, slime, etc. in English. Both morpheme 235 and 236 may be treated as hypermorphemes. [Cf. K. L. Pike, Language, in Relation to a Unified Theory of the Structure of Human Behaviour (Glendale, 1955), Part II, p. 71] That is, morpheme 235 may be analyzed as 215 -s plus -kia?, and morpheme 236 as 215 -s plus -kia? plus 213 -aya plus phonological reduction of /?/ + /a/. This latter analysis would change the relative order of morpheme 213 to the third order, unless -kia? were analysed as a now nonproductive root only occurring in these forms.

[12] The verb -wi- *come* is an exception to this statement. Though -wi- occurs with morpheme 233 -ki, it is not recorded with morpheme 223 -ko.

[13] Morpheme 234 -ti is frequently used and translated in the sense of an action in the future (Cf. Spanish use of 'voy a . . .' with the infinitive). Notice that even the English verb 'to go' with another verb implies a future action: I'm going to hit you. In these cases, as with the Nahuat usage also, the idea of movement is subordinated to the idea of future time.

[14] Herman Aschmann and William Wonderly "Affixes and Implicit Categories in Totonac Verb Inflection" IJAL 18: 130–145 (1952).

in the other .3 allomorphs since the number component is not present in any event. Examples: nasi *I arrive* (< -asi- *arrive*), nikita? *I saw it* (< -ita- *see*); amasikeh *you* (*pl.*) *arrived*, amkitakeh *you* (*pl.*) *saw it*.

Morpheme 241 -h ∞ -keh is a second relative order suffix occurring with all declarative verbs. It may be preceded by a morpheme of Class 210 or of Class 220–30. It may be followed by enclitics only (see below). Morpheme 241 is mutually exclusive with morpheme 242. The distribution of the allomorphs of morpheme 241 is as follows: Allomorph 241.1 -h occurs with the present tense forms of Class I–IV verbs and with two roots of Class V verbs (wi *come* and ia *go*), and with the incompletive past tense and the durative past tense forms of all verbs: weckah *they laugh*, asih *they arrive*, nikawah *we leave it*, patanih *they fly*, tikoniah *we drink it*, niwiceh *we come*, niah *we go*, weckayah *they were laughing*, asiah *they were arriving*, tikoniayah *we were drinking it*, niwiciah *we were coming*, yayah *they were going*. Allomorph 241.2 -keh occurs with the present tense forms of two other Class V verbs, and with punctiliar past and future tense forms of all verbs: nonokeh *we are*, tikpiakeh *we have it*, amkipiakeh *you* (*pl.*) *had it*, kitahtanikeh *they asked for it*, nikmanakeh *we cooked it*, tiktanihkeh *we earned it*; weckaskeh *they will laugh*, nikči·waskeh *we will make it*, pataniskeh *they will fly*, koniskeh *they will drink it*. When allomorph 241.1 -h occurs following 211.1 -?, the combination /?/ + /h/ reduces to /h/: tikonia? + -h > tikoniah *we drink it*, niktania? + -h > niktaniah *we earn it*.

Morpheme 242 -kan 'plural' is a first relative order suffix occurring with all imperative verbs and all permissive verbs. It is immediately preceded by the stem and may be followed by only enclitics. It is mutually exclusive with morpheme 241: šikči·wakan *make it!*, *do it!*, šiakan *go!*, šitakʷa·kan *eat!*, manikči·wakan *let us do it*, matiakan *let's go*, matakʷa·kan *let them eat*.

3.2. Nouns are treated in regard to (1) Root classes, (2) Stem formatives, (3) Possession, (4) Number, (5) Diminutive, (6) Vocative, (7) Locative.

3.2.1. The following summary of noun expressions is given here as a preview of the Nahuat noun before the more detailed distributional discussion which follows. There are two types of minimum-maximum noun expressions—one for Class IA nouns and one for all other classes of nouns.

The minimum noun expression involving Class IA noun roots (minNxA) is composed of a noun root (N) plus a .2 allomorph of the Class 100 morphemes, i.e. 'noun possessor' prefixes. This is expressed in formula as follows: minNxA = +100.2 + N.

The minimum noun expression involving all other classes of noun roots (minNxZ) is composed of simply a noun root, where such roots may occur free, or a noun root plus a stem formative (-t or -ti), i.e. a noun stem, where the noun roots may not occur free.[15] This is expressed in formula as follows: minNxZ = +N/(N + -t/-ti).

The maximum noun expression involving Class IA noun roots (maxNxA) may be composed of a noun stem (Ns)—which may be any stem described below in **4.2**, occurring with the same distribution as Class IA roots—plus a .2 allomorph of Class 100 morphemes, the 'noun possessor' prefixes plus or minus any of the following morphemes, all mutually exclusive with each other: allomorph 22.1 -meh 'simple plural', morpheme 23 -waˑn 'complex plural', morpheme 31 -ciˑn 'diminutive singular', morpheme 32 initial reduplication in pattern C_1V_1- (vowel always short) of suffix morpheme 31 'diminutive plural' (see **3.2.5**),

[15] Free noun roots are roots of Class II (A & B); bound roots are roots of Class I. Class IB roots may occur in either type minimum-maximum expression, i.e. they may occur with .2 allomorphs of Class 100 morphemes or with stem formatives -t or -ti: sin- *corn*: nosin *my corn*: sinti *corn*, gak- *sandal*: nogak *my sandal*: gakti *sandal*.

morpheme 41 -w 'vocative', morpheme 51 -ko 'locative' or morpheme 52 -yo· 'locative'. In addition to the above mutually exclusive constituents of the maximum noun expression for Class IA roots one or two proclitics and up to three enclitics (-wa? *indeed*, -sam- *just, only*, and -ki 'quotative') may occur. This is expressed in formula as follows:

maxNxA = ± procl. +100.2 + Ns ± {22.1, 23, 31, 31 + 32, 41, 51, 52} ± encl.

The pair of braces indicates the mutually exclusive distribution of the morphemic units represented by the enclosed numbers.

The maximum noun expression involving all other classes of noun roots (maxNxZ) may be composed of a noun stem (Ns)—which may be any stem described below (see **4.2**) not occurring with the same distribution as Class IA noun roots—plus or minus any of the following: a .2 allomorph of the Class 100 morphemes, 'noun posessor' prefixes; any allomorph of morpheme 22 (whether prefix or suffix) -meh ∞ reduplication type 1 ∞ reduplication type 2 (see **3.2.4**) ∞ -keh 'simple plural'; morpheme 23 -wa·n 'complex plural'; morpheme 31 -ci·n 'diminutive singular'; morpheme 32 initial reduplication C_1V_1- (vowel always short) of suffix morpheme 31 'diminutive plural' (see **3.2.5**); morpheme 41 -w 'vocative'; morpheme 51 -ko 'locative'; or morpheme 52 -yo· 'locative'. In addition to the above mutually exclusive constituents of the maximum noun expression for non-IA class roots one or two proclitics and up to four enclitics (-ti 'stem formative', -wa? *indeed*, -sam- *just, only*, and -ki 'quotative') may occur. This is expressed in formula as follows:

maxNxZ = ± procl. ± 100.2 ± {22.2, 22.3 . . . +Ns . . . 22.1, 22.4, 23, 31, 31 + 32, 41, 51, 52} ± encl.

The pair of braces indicates the mutually exclusive quality of the morphemic units represented by the numbers enclosed therein; the numbers preceding . . . +Ns are prefixes and those following +Ns . . . are suffixes. The three dots surrounding +Ns are meant to indicate that a selection is made of morphemic material within the braces such that a possible sequence may include any one affix and Ns.

3.2.2. The occurrence of noun roots with or without allomorphic alternants determines two main classes of nouns. Class I occurs without allomorphic alternants and is subdivided by the distribution of these roots with the .2 allomorphs of Class 100 morphemes (i.e. allomorphs 113.2 no-, 121.2 to-, 122.2 amo- and 133.2 i-∼iy-) and the occurrence of certain stem formatives. Class II occurs with allomorphic alternants and is subdivided by the type of allomorphic alternants which occur.

The main noun classes and their subclasses are here described.

Class I (without allomorphic alternants):

Subclass A—Roots occurring obligatorily with the .2 allomorphs of Class 100 morphemes: -ya? *nose*, -ma· *hand*, (and most other body parts), -pih *aunt*, -coyo *younger sibling*, (and all other kinship terms), -pan *place*, (and all other locative roots), -čam *home*.

Subclass B—Roots occurring optionally either with the .2 allomorphs of Class 100 morphemes or with a formative of the free stem (either -ti or -t). The stem formative -ti and the prefixes constituting the .2 allomorphs of Class 100 morphemes are mutually exclusive in occurrence with noun roots, but -t is not, i.e. .2 allomorphs of Class 100 morphemes and stem formative -t may occur simultaneously in the same noun construction. Roots occurring with -ti: oh- *road*, es-[16] *blood*, sin- *corn*, tapan- *house beam*, čakis- *gnat*, miš- *cloud*, gak- *sandal*,

[16] The root es- *blood* may occur also with morpheme 323 -yo 'noun stem formative' instead of with -ti. The occurrence of morpheme 323 -yo with es- *blood* is obligatory when any .2 allomorph of a Class 100 morpheme occurs.

kʷač- *shirt*, nakas- *ear*.[17] Roots occurring with -t: o·mi- *bone*, taga- *man*, kone- *child*, šiwi- *leaf*, šolo- *baby*, wewe- *old man*, siwa- *woman*, a·ma- *paper*, saka- *grass*, ohta- *bamboo*, capo- *banana*, ci·ka- *ant*, tiawa- *rain*, etc.

Subclass C—Roots occurring as free forms and optionally with the .2 allomorphs of the Class 100 morphemes: ahayo *beans*, peloh *dog*, piyo· *chicken*, ko·yame *pig*, šikal *gourd basin*, mačete *machete*, cocol *cloth*, taškal *tortilla*.[18]

Class II (with allomorphic alternants):

Subclass A—Roots occurring as free forms of the shape CVhli, but with bound stem allomorphs of the shape CVl (i.e. šahli *sand* alternates as a free form with the bound form -šal) and optionally with the .2 allomorphs of the Class 100 morphemes: šahli *sand*, tahli *earth*, kahli *house*, tihli *soot*, mihli *cornfield*.

Subclass B—Miscellaneous free roots with one bound allomorph occurring optionally with the .2 allomorphs of the Class 100 morphemes or as a bound alternant in polymorphemic constructions (i.e. occurrences of affixes or in compound with other roots or both): yolo ∞ -yol- *heart*, kʷawit ∞ -kʷa- *wood*, kʷeyi ∞ -kʷe- *skirt*.

3.2.3. The phonemes /o/ plus /i/ reduce to /o/ in the combination noun-possessor plus -ikni *sibling*: no- + -ikni > nokni *my sibling*. A combination of short vowel plus long vowel does not reduce, nor do other combinations of vowels reduce in the noun constructions: noa·t *my water*, moo·miyo *my bone*, noyi·š *my eye*,[19] noohwi *my road*, toohtat *our bamboo*.

The possession indicators of the nouns are the corresponding allomorphs of the person morphemes, i.e. the .2 allomorphs of Class 100 morphemes: 113.2 no-, 120.2 mo-, 121.2 to-, 122.2 amo- and 133.2 i- ~ iy. They are second relative order prefixes which may be followed only by allomorphs 22.2 and 22.3 of morpheme 22 'simple plural'. They are obligatory for Class IA noun roots and optional for all others. For examples and complete description see **3.1.3.**

3.2.4. There are two number morphemes for nouns: morpheme 22 'simple plural' and morpheme 23 'complex plural'. All number morphemes and their allomorphs are mutually exclusive with each other and with all other noun suffixes.

Morpheme 22 -meh ∞ reduplication type ✻1 ∞ reduplication type ✻2 ∞ -keh 'simple plural' occurs optionally with all noun roots. The distribution of the allomorphs of morpheme 22 is as follows. Allomorph 22.1 -meh occurs with the large majority of noun roots; it occurs with some members of every root class,[20] and always immediately follows the root or stem: nopihmeh *my aunts*, ičameh *their homes*, ipiyo·meh *her chickens*, moko-yamemeh *your pigs*, notaškalmeh *my tortillas*, kahlimeh *houses*, mišmeh *clouds*, šolomeh *babies*, a·mameh *papers*. Allomorph 22.2 reduplication type ✻1 which is initial reduplication of pattern C_1V_1- of the root or stem is a first order prefix which occurs with the following noun roots only: kone- *child*, nakas- *ear*. Examples: kokonet *children*, inanakas *his ears*. Allomorph 22.3 reduplication type ✻2 which is initial reduplication of pattern C_1V_1h- of the root or stem is a first order prefix which occurs with the following noun root only: taga- *man*, tahtagat *men*. Allomorph 22.4 -keh is a suffix occurring immediately following these noun stems only:[21] siwat *woman*, wewet

[17] Certain roots of this class occur with an alternant form in derived verb stems (see **4.1**).

[18] Possibly all /l/ final nouns belong to this class.

[19] With the long vowel root -i·š *eye* the forms noy- *my*, moy- *your*, iy- *his* (*etc.*), toy- *our*, amoy- *your* (*pl.*) occur.

[20] Certain forms may occur with either of two plural markers: tagameh *men*:tahtagat *men*; siwameh *women*:siwatkeh *women*. The latter form in each case is the more frequently heard. Note, however, that only -pilowa·n *children* occurs.

[21] The forms siwa- *woman*, wewe- *old man* and ipilci·n *child* in the plural constructions do not

old man. Examples: siwatkeh *women*, wewetkeh *old men*.

Morpheme 23 -wa·n 'complex plural', i.e. plurality of objects plus plurality of source occurs optionally with a limited group of kinship terms. It immediately follows the root or stem and is mutually exclusive with morpheme 22 and all other suffixes. Its distribution is limited to kinship terms relating to a group of persons capable of having more than one mother (or presumably father) for the group as a whole, e.g. cousins, sons and daughters of my father's and mother's brothers and sisters, which would, as a group, have several mothers and fathers. Examples of the members of this group in Nahuat: -coyo *younger sibling*, -koko· *older sibling*, -ikni *sibling in general*, -wes *sister-in-law* (*woman speaking*), -teš *brother-in-law* (*woman speaking*), -mon *son-in-law* -siwamon *daughter-in-law*, etc. Examples of the plural morpheme: nocoyowa·n *my younger siblings* (*various mothers*), ikoko·wa·n *his older siblings* (*various mothers*), moweswa·n *your sisters-in-law* (*various mothers, woman speaking*).

Many nouns never occur with any plural form: mecti *moon*, tahli *earth*, sitalin *star*, etc.[22]

3.2.5. The diminutive is indicated by morpheme 31 -ci·n for the singular of all nouns. It is optional in occurrence and immediately follows the root or stem when singular. It may be preceded by morpheme 32 reduplication type #3 'diminutive plural'. Examples: peloci·n *little dog*, siwaci·n *little woman*, tagaci·n *little man*, kalci·n *little house*.

The plural of the diminutive forms is indicated by morpheme 32 initial reduplication of type C_1V_1- (with vowel always short) of the diminutive suffix morpheme 31 -ci·n. Morpheme 32 is, therefore, first relative order mutually obligatory with morpheme 31 in plural constructions. Examples: pelocici·n *little dogs*, siwacici·n *little women*, tagacici·n *little men*, kalcici·n *little houses*.

Morpheme 31 is mutually exclusive with all other suffixes and morpheme 32 is mutually exclusive with all other noun plural morphemes as well.

3.2.6. The vocative is indicated by morpheme 41 -w for all nouns and occurs immediately following the stem. It is mutually exclusive with all other suffixes. This construction is statistically quite rare. Morpheme 41 always immediately follows the stem and is never followed by any morpheme except an enclitic. Examples: taow *girl!*, čow *boy!*.

3.2.7. The locative is indicated by two suffixes: morpheme 51 -ko and morpheme 52 -yo·. Both are first relative order suffixes and mutually exclusive with all other noun suffixes and with each other. They may be followed only by enclitics. The locative is also indicated by compounds of a noun root plus a locative root of Class IA. For a description of these compounds see **4.2.**

The semantic distinctions indicated by the locative suffixes have not yet been precisely determined, but the following approximations may be stated: -ko defines location in, and -yo· defines a general location of abundance. Examples: itenko *in his mouth*, iyihtiko *in his stomach*, tiko *in the fire*, ahkoyo· *bamboo attic*, akawalyo· *place of jungle-type growth*, talpinolyo· *dusty place*, teyo· *stony place*, tepeyo· *hill*.

3.3. Pronouns are treated in regard to roots, number and emphasis-respect.

3.3.1. Pronoun roots occur as the .1

yield a satisfactory analysis as yet. Both siwa- and wewe- seemingly retain the -t stem formative in the most commonly heard forms (fn. 20 above) whereas no other noun roots do. The plural of -pilci·n *child* is -pilowa·n *children* where an unexplained /o/ occurs. The singular form -pilci·n never occurs without the diminutive morpheme, #31, -ci·n.

[22] Many nouns occur with plural meaning when juxtaposed to a numeral or other qualifying word: o·me ko·yame *two pigs*, eyi peloh *three dogs*, nawi tagat *four men*, miak mihli *many cornfields*.

allomorphs of the following four morphemes: 113, 121, 122, and 133, i.e. 113.1 neh 'first person', 121.1 teh 'second person inclusive', 122.1 ameh- 'second person plural' and 133.1 yeh 'third person'. Three of the four morphemes occur as singular free forms as follows: neh 'first person singular' *I/me*, teh 'second person singular' *you*, yeh 'third person singular' *he/him*, etc. Combinations of these four morphemes with a plural morpheme form plural pronoun constructions (see **3.3.2**).

The third person singular personal free form, yeh, serves also as a relative pronoun, *that, who, which*, as in inón tagat yeh ne·pa *that man who is there*, and inón komit yeh yeh kičí·h *that pot that she made*.

3.3.2. The number morphemes indicate two types of plural in pronouns. Morpheme 62 -wa·n ∞ -amen indicates the specific plural as opposed to morpheme 63 which indicates the representative or generic plural. Morpheme 62 is a suffix of first relative order from the stem and may be followed only by morpheme 64 -wa 'emphasis' or enclitics. It is mutually exclusive with morpheme 63 'generic plural'. The distribution of morpheme 62 with its allomorphs is as follows: Allomorph 62.1 -wa·n occurs only with 122.1 ameh- 'second person plural': amehwa·n *you plural*. Allomorph 62.2 -amen occurs with all other pronoun roots: nehamen *we/us excl.*, tehamen *we/us incl.*, yehamen *they/them*.

Morpheme 63 -emen indicates the representative or generic plural for all pronouns. It is a suffix of the first relative order from the stem and may be followed only by morpheme 64 or enclitics. It is mutually exclusive with morpheme 62 'specific plural'. Examples: nehemen *we/us excl.*, tehemen *we/us incl.*, amehemen *you plural*, yehemen *they/them*.

The meaning contrast between morphemes 62 and 63 is more fully described and illustrated as follows: Morpheme 62 indicates a delimited group usually present in its entirety, whereas morpheme 63 indicates a group in general, large and only generally defined, with at least one representative of the group present. Examples: amehwa·n anyahkih yalwa *you (plural) went yesterday*, amehemen antahtowah inglés *you (plural) talk English*,[23] nehamen nitahtowah mehikano *we (excl.) talk Aztec*, nehemen nitahtowah mehikano *we (excl.) talk Aztec*, yehamen takʷikah *they sing*, yehemen takʷikah *they sing*.

3.3.3. Morpheme 64 -wa indicates emphasis or respect with the singular forms of the pronouns.[24] Morpheme 64 is first relative order and mutually exclusive with all other noun suffixes. Examples: nehwa *I/me*, tehwa *you (sg.)*, yehwa *he/him, etc.*

3.4. Particles are roots which never occur inflected. There are two classes of particles in Nahuat: free particles and bound particles. Bound particles are divided into two classes as follows: proclitics are bound particles which occur preceding the verb root and its affixes, which when combined with a verb construction do not shift the stress, and which in some cases occur with more than one word class; enclitics are bound particles which occur following the word, which when combined with a verb construction do not shift the stress, and which in some cases occur with more than one word class.

3.4.1. Free particles are single morphemes

[23] The first two examples were volunteered to me by Sr. Juan Salas, Mecayapan, Ver., but he also stated that he could make no distinction between the uses of the two pluralizing morphemes with other than the second person. The author suggests that with other than the second person the distinctions are too abstract and subtile for the native speaker to note and verbalize. Other informants could not even make a clear statement of the distinctions in the occurrence of the forms, although all recognized a contrast in meaning.

[24] It is possible that the plural forms occur inflected for emphasis, too. Note the forms nehamen-wa? *we, indeed* and amehwa·n-wa? *you (pl.), indeed* which have been recorded. It may be that the form -wa? described in **3.4** as an enclitic is identical with the form -wa described above as a pronoun emphasizer. (See also fn. 26.)

such as te· *what?* kan *where*, inin *this*,[25] seh *again*. Nahuat numerals are also included in this class and are regularly used from 'one' to 'four': se *one*, o·me *two*, eyi *three*, nawi *four*. Spanish numerals are used from 'five' on without phonetic or morphological adaptation.

3.4.2. The bound particles are seven in number, one of which always precedes the word to which it is attached, five of which always follow the word to which they are attached, and one of which may either follow or precede (without perceptible change in meaning) the word to which it is attached. They are described as follows:

iš- always precedes the word construction (a member of any word class may occur), may be followed by -sam- (see below) and indicates an interrogative type construction which demands an affirmative answer, a negative answer or an evasive answer such as 'perhaps', 'who knows', etc. Examples: iš-tia *Are you going?*, iš-tikpia? *Do you have it?*, iš-até· *Isn't there any?*.

-ha ~ -ya is a first order enclitic, i.e. always precedes all other enclitics and immediately follows the stem. It occurs only with verbs. Its function is similar to (and possibly related to by borrowing) the Spanish particle 'ya' *already, now*. The distribution of -ha ~ -ya is: -ha occurs after /h/ and all vowels; -ya occurs elsewhere. Examples: kičihkeh-ha *they've already done it*, yahkih-ha *he's already gone*, kikʷah-ha *he's already eaten it*, nikowako-ha *I already came and bought it*, kipiaya-ha *he already had it*, nikmagaya-ha *I was already hitting him*, kisa?-ya *he's already left*, asi?-ya *he's already arrived*, nikmaga?-ya *I already hit him*.

-wa? is a second order enclitic which may be preceded by either -ya or -ti and possibly others. It occurs with all word classes. Its function is similar to that of morpheme 64 -wa 'pronoun emphasis-respect' (see **3.3.3**).[26] Examples: tami?-wa? *all gone, indeed!*, kičihkeh-wa? *they did it, indeed!*, šia-wa? *go, indeed!*, kisa?-yawa? *indeed, he's already gone!*, teh-wa? *you, indeed!*, aya-wa? *not yet, indeed!*, taškal-wa? *tortilla, indeed!*.

-o? *just barely* is postposed to the entire verb construction plus any other enclitics occurring except -ki (see below). Its exact relative order among the enclitics is at present indeterminate due to the poverty of recorded occurrences. Examples: tami?-o? *it's just barely finished*, nikčih-o? *I just barely did it*, yahkih-o?ki *it is said that he just barely went*.

-ki 'quotative' is terminal to the entire verb construction or any word class with which it occurs and all enclitics that may occur with the construction. Examples: yahkih-ki *she said, he's gone*, awél-ki *he can't, he said*, ayo-ki *no more, he said*, taškal-ki *tortillas, he said*, yehamen-ki *they, it is said*.

-sam- *just, only* is a first order proclitic and indeterminate order enclitic, i.e. it occurs immediately preceding the stem, but the order following the stem is not at present determined. Examples: išsam-kičih *Did he just finish doing it?*, sam-yahkih *he just went*, kičih-sam *he just did it*, yahkih-sam *he just went*, kita?-sam *he only saw it*.

4. Under the following section on derivation are discussed morpheme sequences which form stems. These stems are classified into the four word classes described above (in **3**) by the same criteria discussed in that section. The derived stems are discussed according to the word class of the resultant stem regardless of the constituent elements occurring.

[25] It is possible to consider inín a reduplicated form of *in. The form *in only occurs freely in rapid speech, and imín contrasts with one word only: inón *that*. However, notice the recurring partials in the following words: imanín (freely alternating with in inanín) *quickly*, inám *now*. Such an analysis appears comparable to a sub-analysis of th in this, that, those, these, etc., in English.

[26] In slow or precise speech, particularly from old women, -wa? becomes -wač as in: kena-wač *yes, indeed*, iga-wač 'emphatic agreement'.

Derived stems are of two types: Root plus root, and root plus stem formative or derivative affix. Derived stems may be composed of two or more roots; stem formatives may occur with a root or with a derived stem of two or more roots.

4.1. Verb derivation is discussed as to stem types and formation and then a formulaic summary of the verb derivation is presented.

4.1.1. In the root plus root constructions the second consonant of a noun stem when voiceless is replaced by its voiced counterpart in the derived verb stem: -nakas *ear* + -witeki- *beat* > -nagaswiteki- *ear-ring* (cf. the same phenomenon in derived noun stems **4.2**).

Only one type of root plus root construction occurs, i.e. a noun root or stem plus a verb root or stem, or a unique[27] constituent such as -con-, -teh-, -pos-, -sen- plus a verb root or stem.[28] Examples: -ma·posteki- *break an arm, hand,* < -ma· *hand, arm* + -posteki- *break*; -taltemi- *earthfill* < -tal- *earth* + -temi- *fill*; -conteki- *cut with a machete* < -con- + -teki- *cut*; -yokači·wa- *do alone* < -yoka- *aloneness* + -či·wa- *do, make*; -senehnemi- *walk together* < -sen- + -ihli- *say something together to someone* < -sen- + ihli- *say something to someone.*

Various types of root plus stem formative affix derivations occur and are discussed below. The distribution of the stem formatives is in terms of occurrence within the stem, i.e. inflectional affixes may occur juxtaposed to the stem formatives discussed below, but their distribution is in terms of inflectional affixes to the stem.

Type 1. A Spanish loan verb infinitive plus morpheme 317 -o- 'stem formative', suffix of the first relative order from the root. It may be followed by morpheme 311, 313, 312, 314 or 215. Stems of this type occur with the same inflectional morphemes as do Class II verb roots. Examples: -hugar- *play* + -o-: nihugarowa *I play*; -kontar- *count* + -o-: nikontarowa *I count it.*

Type 2. A Class II verb root, or a Spanish loan stem (i.e. Type 1 above), plus morpheme 311 -wa ∞ -wi 'stem formative in the present tense', suffix of the second relative order from the root. It may be preceded only by morpheme 317 -o- and may be followed by morpheme 313, 312, 314 or 215. Stems of this type occur with the same inflectional morphemes as do Class II verb roots. Examples: -či·- *do, make* + -wa: nikči·wa *I do it, I make it*; -ka- *remain, leave* + -wa: nikawa *I leave it,* but -ka- *remain, leave* + -wi: nikawi *I remain*; -hugar- *play* + -o- + -wa: nihugarowa *I play.*

Type 3. Three initial reduplicatives of first relative order from the root occur as stem formatives. They are as follows: Morpheme 318 C_1V_1-[29] indicates a derived but non-predictable meaning from the root. Examples: -noca- *call*: -nonoca- *counsel, warn, advise*; -powa- *cleanse, wipe*: -popowa- *relate what happened*; but also -kalaki- *enter*: -kakalaki- *enter some house many times.*[30] Morpheme 319 C_1V_1h- indicates simple repeated action. Examples: -noca- *call*: -nohnoca- *fight verbally*; -pano- *pass by*: -pahpano- *pass by many times*; -kalaki-

[27] A unique constituent is an otherwise non-occurring morpheme.

[28] The phrase root or stem is used to allow for the description of the successive layering of formations in Nahuat derivation. A stem may include either a simple root or a previously derived stem, etc.

[29] Note that with some stems Reduplication types #1 and #2 (C_1V_1 and C_1V_1h) may be semantically equal. However, some of these stems may also have freely alternating sememes: -kahkalaki *enter same house many times* or *many people enter one house at the same time*, -posteki *break*: -popostekí *break many times* or *break in many pieces.* The form -pohposteki also means *break many times* or *break in many pieces.* The forms -teki- *cut* and -maka- *hit* function as does -posteki, but the form -paka- *wash* has -papaka- *wash many times,* but -pahpaka *wash many things.*

[30] See footnote 29.

enter: -kahkalaki- *enter same house many times.*[31] Morpheme 320 C_1V_1·- indicates repeated action with reference to various objects. Examples: -kalaki- *enter*: -ka·kalaki *enter various houses one after another*; -kʷepa- *turn over or up*: -kʷe·kʷepa- *turn over or up many objects.*

Type 4. A verb root or stem of types 1–3 plus morpheme 313 -li 'second object', with qualitative change of final root or stem vowel to /i/. Morpheme 313 is a suffix of the third relative order from the root. It may be preceded by morpheme 317 and 311, and it may be followed by morpheme 313, 314 and 215. Stems of this type occur with the same inflectional morphemes as do Class IV verb roots. Examples: -či·wili- *make, do something for someone else* < -či·- *make, do*; -kawili- *leave something for someone else* < -ka- *leave.*

Type 5. A noun root, a noun stem of two noun roots, a verb root stem of Type 1–4 above, plus morpheme 312 -ti 'causative stem formative', suffix of the fourth relative order from the root. It may be preceded by morpheme 315, 317, 311, and 313; it may be followed by morpheme 314 and 215. The actual occurrence of morpheme 312 is limited to a somewhat infrequent occurrence by the following considerations:

(1) A construction of the form -teki- with certain verb roots indicates obligation and frequently substitutes for the causative construction, e.g. niktekimaga iga kičiwa *I 'made' him do it,* neˑtekimagakeh iga nikwiga *they 'made' me bring it.*

(2) The phonetic shape of morpheme 312 is identical with that of morpheme 234 -ti. Distributionally both 312 and 234 may occur in the same positions, i.e. immediately following the root, resulting in potential ambiguity. Hence, a paraphrastic syntactic construction may also be used instead of the more ambiguous causative construction. Examples: ničokati *I'm going to cry* or *I cause him to cry*, nikmagati *I'm going to hit him* or *I cause him to be hit*, etc. The syntactic construction is: nikčiwa iga makimaga *I cause him to be hit, I cause him to hit it*, nikčiwa iga mačoka *I cause him to cry.*

With morpheme 312 the stem final vowel of certain roots is lost and the stem-final consonant palatalizes according to the following pattern:[32] s > š, č > ch, t > č. Examples: -kišti- *take out* < -kisa- *come out*, -nešti- *show* < -nesi- *appear*, -pahti- *heal* < -pačiwi- *get well*, -mačti- *teach* < -mati- *know.*

Causative stems occur with the same affixes as roots of Class IV. Examples: -yeti- *get another mother* < -ye- *mother*, -močanti- *make one's home* < -čam *home*, -čokati- *cause to cry* < -čoka- *cry*, -tehkawilti- *put something upon something else higher* < -tehkawi- *ascend*, -kʷakolalti- *make a wooden fence* < -kʷa- *wood* plus kolal (from Spanish corral) *enclosure.*

Type 6. A stem of Type 5 above plus any of the following verb roots: -nemi- *foot-move*, -wi- *come*, -ia- *go*, -ka- *remain.* Stems of this type select the same suffixes as do the final constituents of the stem. Examples: -kočtika- *remain sleeping* < -koči- *sleep* plus -ti stem formative (see type 5 above) plus -ka, -kickiti- *go holding on to something* < -kicki- *grasp* plus -ti stem formative (see type 5 above) plus -ia- *go*, -patanti- *go flying* < -patani- *fly* plus -ti stem formative (see type 5 above) plus -ia- *go*, -kočtinemi- *sleep walk* < -koči- *sleep* plus -ti stem formative (see type 5 above) plus -nemi- *walk.*

Type 7. A verb root of any class or any previous stem plus morpheme 314 -to 'stative indicator' suffix of the fifth relative order from the root. It may be preceded by

[31] See footnote 29.

[32] The evidence from one noun suggests treatment of the phenomenon as an automatic type of change. Note tišti *ground corn dough* < -tisi- *grind* plus -ti stem formative for nouns.

morpheme 315, 317, 311, 313, or 312; it may be followed by 215. Stems of this type occur with the same inflectional morphemes as does -ono- *be* of Class V. Examples: -siato- *be tired* (cf. -siawi- *be tired* a Class I root), -wakto- *be dry* < -waki *dry*, -mikto- *be dead* < -miki- *die*, -to·kto- *be planted, buried* < -to·ka- *plant, bury*.

Type 8. A verb root or any previously described stem plus morpheme 215 -s 'future tense' suffix of the sixth relative order plus verb root -neki- *want*. Morpheme 215 may be preceded by morpheme 315, 314, 312, 313, 311 or 317. It may be followed by only another verb root or stem. Stems of this type occur with the same inflectional morphemes as occur with the Class I root -neki- *want*. Examples: -mikisneki- *want to die* < -miki- die, -tahtosneki- *want to talk* < -tahto- *talk*, -yawisneki- *want to go* < -ia- *go*.

Type 9. A noun stem plus morpheme 315 -ta 'verbalizer *# 1'* suffix of the first relative order from the root. (See below types 11 and 12 for verbalizer *# 2.*) Morpheme 315 may be followed by morpheme 312, 314, or 215. Stems of this type occur with the same inflectional morphemes as Class III roots. Examples: -sokita- *be muddy* < sokit *mud*, -ehekata- *be windy* < ehekat *wind*, -mangota- *go mango-gathering* < mango *mango*, -šoteta- *go snail-gathering* < šote *snail*.

Type 10. A noun (root, stem or possessed form) plus a .3 allomorph of Class 100 morphemes 'verb subject' prefix of the second relative order from the root. It may be preceded by modal prefix 11 a- 'negative'. Examples: nitagat *I'm a man* < tagat *man*, niweweh *I'm an old man* < weweh *old man*, anisiwat *I'm not a woman* < siwat *woman*, iš-tiburroh *Are you a donkey?* < burroh *donkey* (from Spanish), tiipilci·n *you are his son* < ipilci·n *his son*.

Type 11. The root of a Class I Particle stem (see below under **4.3**) plus 311.1 -wa 'transitivizer'[33] plus 316 -ya 'verbalizer *# 2'* suffix of the second relative order from the root. Morpheme 316 may be preceded by only 311.1; it may be followed by morpheme 312, 314 or 215. Stems of this type occur with the same inflectional morphemes as Class III roots. Examples: -tomawaya- *become fat* < tomakti *fat*, -tak^{w}awaya- *become hard* < tak^{w}akti *hard*.

Type 12. The root of a Class II Particle stem (see below under **4.3**) plus morpheme 316 -ya (see above). Stems of this type occur with the same inflectional morphemes as Class III roots. Examples: -etiya- *become heavy* < etik *heavy*, -šokoya- *become bitter* < šokok *bitter*, -istaya- *become white* < istak *white*.

Type 13. The stems of types 9, 10, 11 and 12 may each occur with morpheme 312. Stems of type 13 occur with the same inflectional morphemes as Class IV roots.

Type 14. Morpheme 145 ta- plus a verb root. This type stem occurs only in noun stem derivation (see **4.2.1**). Examples: ta- + -k^{w}a·- *eat* + -l 'nominalizer' > tak^{w}a·l *meal*, ta- + -ihto·- *speak* + -l 'nominalizer' > tahto·l *word*.

4.1.2. Below is summarized the types of verb stem derivation just described and here expressed in formulas. The following symbols are used: N = noun root or stem, V = verb root or stem, V_{II} = Class II verb root, V_s = verb stem, Sp. = Spanish verb infinitive, Pt_{II} = root of Class II Particle stem.

N + V = V_s; Sp. + 317 + 311 = V_s; V_{II}/Sp. + 311 = V_s; reduplicative + V = V_s; any preceding V + 313 = V_s; any preceding V or N + 312 = V_s; any preceding V or N + 312 + -nemi-, -wi-, -ia-, or -ka-

[33] Allomorph 311.1 -wa functions as a transitivizer only when it is in alternation with allomorph 311.2 -wi in the same grammatical series. When 311.1 does not appear in complementary distribution with 311.2, allomorph 311.1 functions as a simple stem formative. Only 311.1 occurs with Spanish loan words.

= V_s; any preceding V + 314 = V_s; any preceding V + 215 + -neki- = V_s[34]; N + 315 = V_s; Class 100.3 + N = V_s; Pt_I + 311 + 316 = V_s; Pt_{II} + 316 = V_s; (N + 315), (Class 100.3 + N), (Pt_I + 311 + 316), or (Pt_I + 311 + 316) + 312 = V_s.

4.2. Noun derivation is discussed as to root plus non-root types and root plus root types.

4.2.1. There are three types of root or stem plus non-root derivative constructions which occur: (1) noun root plus obligatory stem formative of the type referred to in **3.2.2**, (2) noun plus derivative affix, and (3) verb plus derivative affix.

Four stem formatives occur. Morpheme 321 -ti occurs with Class IB noun roots. Examples: ohti *road*, esti *blood*, sinti *corn*. Morpheme 322 -t occurs with Class IB noun roots and is a first order suffix. It may be followed only by enclitics.[35] It is mutually exclusive with all other noun suffixes. Examples: o·mit *bone*, tagat *man*, šiwit *leaf*, etc. Morpheme 323 -yo occurs with certain words of Class IC, i.e. physiological terms such as 'flesh', 'blood', etc. A .2 allomorph of Class 100 morphemes is an obligatory element to a construction with stem formative -yo: nonakayo *my flesh*, inakayo *his flesh*, moesyo *your blood*, iyesyo *his blood*, iyehekayo *her breath*, toehekayo *our breath*. Morpheme 324 -wi occurs with only oh- *road*. It is also in mutually obligatory relation with the .2 allomorphs of Class 100 morphemes in that if -wi occurs a .2 allomorph must also occur: noohwi *my road*, nok^{w}aohwi *my hair-part*.

The following derivative affixes occur with nouns: Noun root or stem plus morpheme 331 -h 'collective indicator', noun root or stem plus morpheme 332 pec- 'content simplicity indicator' preposed to the root or stem, or noun root or stem plus morpheme 333 namol- 'composition simplicity indicator' preposed to the root or stem. Examples: k^{w}awih *woods* < k^{w}awit *wood*, pectaškal *plain tortilla* < taškal *tortilla*, peca·t *plain water* < a·t *water*, ipeckan *his nakedness* < -kan *place*, namolsombrer *plain hat* < sombrer *hat* (from Spanish), namolk^{w}awit *soft, non-durable wood* < k^{w}awit *wood*.

The following derivative affixes occur with verbs: Verb root or stem plus morpheme 334 -lo 'pseudo-passive indicator', verb root or stem plus one of the following nominalizing suffixes: morpheme 335 -l, morpheme 336 -ni or morpheme 337 -lis. Examples: pewalo *beginning* < -pewa- *begin*, nemilo *a certain type of communal gathering* < -nemi- *foot-move*, mikilo *activity concerning dead people* < -miki- *die*, tak^{w}a·l *meal* < -tak^{w}a·- *eat*, tahto·l *word* < -tahto·- *speak*, te·k^{w}a·ni *tiger* < -te·k^{w}a·- *eat people*, conkalwiwitani *dragon fly* < -conkalwiwita- *hair-pull*, kisalis *diahrrea* < -kisa- *go out*, tamačtílisti *teaching* < -tamačti- *teach*.

When a verb root or stem occurs with morpheme 335 -l 'nominalizing suffix', the second consonant of the root or stem when voiceless is replaced by its voiced counterpart in the derived noun stem: -pepečowa- *mud a house wall* < tabepečol *mud wall*. (Cf. the same phenomenon in derived verb stems, **4.1.**)

4.2.2. The two types of root plus root derivative constructions which occur: (1) Noun root or stem plus noun root or stem and (2) verb root or stem plus a noun root or stem (usually a Class IA locative root). Examples: (N + N) k^{w}etaškoštal *leather bag* < k^{w}etaš *leather* + koštal *bag, sack* (from Sp. costal), piyo·teksis *hen egg* < piyo· *chicken* + teksis *egg*, talpinol *dust* <

[34] At this point there are no new derivational affixes to be added, but morphemes 311, 312, 313 and 314 can be added singly or all at once. The root -neki- *want* usually occurs only once in any given derived stem, but theoretically could occur an infinite number of times following any or all morphemes 311–14.

[35] See **3.4.**

tal- *earth* + pinol *pinole*, tiopan *church* < tio *God* + -pan *place*; (V + N) tako·mičkan *dark place* < tako·mik *it is dark* + -kan *place*.

4.3. Three types of derived particle stems occur all of which are root plus derivative affix. Type 1 consists of root plus -kti: šošokti *green*, tomakti *fat*, picakti *thin and narrow*, kanakti *thin and flat*, šolakti *slippery*.

Type 2 consists of root plus -k: pistik *black*, etik *heavy*, yektik *good, right*, čičik *bitter*, copelik *sweet*.

Type 3 consists of forms derived from morpheme 12 a- 'negative' plus roots such as te· *what?*, kan *where, place*: até· *nothing*, nité· *nothing* (with negative constructions), akanah *nowhere*.

Certain forms, which for lack of data are classified as simple particle roots, contain suspicious recurring partials. Notice the similarity in the following: aman *today*, naman *later today*, imanín *quickly*, keman *when?*; o·mpa *there* (*out of sight*), ne·pa (*in sight*), isahpa *early* (*in time of day*); ayo *no more*, aya *not yet*, ayá *no* (with indicative verbs), amo *no* (with imperative and permissive verbs). These latter with /a/ initial could be analyzed as negative verbal prefix, morpheme 12, plus a unique constituent (-yo, -ya and -mo) except for the difficulty that morpheme 12 is always atonic in all other places.

5. The morphemes and their allomorphs of Isthmus Nahuat are alphabetically listed with their meanings as follows:

a- #11 'negative mode prefix'; am- #122.3 'second person exclusive plural verb prefix'; ameh- #122.1 'second person exclusive plural pronoun root'; -amen #62.2 'specific plural suffix of pronouns'; amo- #122.2 'second person exclusive plural noun possessor prefix'; -aya #213 'incompletive past suffix'; -a? #211.2 'present tense suffix'; -a? #214 'durative past suffix'.

-ci·n #31 'diminutive singular'.

-emen #63 'generic plural suffix of pronouns'.

-h #212.2 'punctiliar past suffix'; -h #241.1 'plural suffix of verbs'; -h #331 'noun derivative suffix'.

i- ~ iy- #133.2 'third person noun possessor prefix'.

-ka #221 'complete aspect suffix'; -kan #242 'plural suffix of verbs'; -keh #22.4 'simple plural suffix of nouns'; -keh #241.2 'plural suffix of verbs'; ki- #133.4 'third person verb object prefix'; -ki #222 'complete aspect suffix'; -ki #233 'incomplete aspect suffix'; -ko #223 'complete aspect suffix'; -ko #51 'locative suffix'.

-l #335 'noun derivative suffix'; -li #313 'second object suffix'; -lis #337 'noun derivative suffix'; -lo #334 'noun derivative suffix'.

ma- #12 'permissive mode prefix'; mic- #125.4 'second person indeterminate number verb object prefix'; -meh #22.1 'simple plural suffix of nouns'; mo- #120.2 'second person singular noun possessor prefix'; mo- #146 'reflexive and passive voice prefix'.

namol- #323 'noun derivative prefix'; ne·- #113.4 'first person verb object prefix'; neh #113.1 'first person pronoun root'; ni- #113.3 'first person verb subject prefix'; -ni #336 'noun derivative suffix'; no- #113.2 'first person noun possessor prefix'.

-o- #317 'verb stem formative'.

pec- #322 'noun derivative prefix'.

Reduplication type #1, C_1V_1- #22.2 'simple plural prefix of nouns'; reduplication type #1, C_1V_1- #318 'verb stem formative'; reduplication type #2, C_1V_1h- #22.3 'simple plural prefix of nouns'; reduplication type #2, C_1V_1h- #319 'verb stem formative'; reduplication type #3, C_1V_1 short #32 'diminutive plural'; reduplication type #4, C_1V_1· #320 'verb stem formative'.

ši- #126.3 'second person imperative verb subject prefix'.

-s #215 'future tense suffix'; -skia? #235

'frustrative aspect suffix'; -skiaya ⁂236 'conditioned aspect suffix'.

-t ⁂322 'noun stem formative suffix'; ta- ⁂145 'third person indeterminate number verb object prefix'; -ta ⁂315 'verbalizer suffix'; te·- ⁂121.4 'second person inclusive verb object prefix'; teh ⁂121.1 'second person inclusive pronoun root'; ti- ⁂121.3 'second person inclusive verb subject prefix'; -ti ⁂234 'incomplete aspect suffix'; -ti ⁂312 'causative suffix'; -ti ⁂321 'noun stem formative'; to- ⁂121.2 'second person inclusive noun possessor prefix'; -to ⁂224 'complete aspect suffix'; -to ⁂314 'stative suffix'.

-w ⁂41 'vocative suffix'; -wa ⁂64 'emphasis-respect suffix'; -wa ⁂311.1 'verb stem formative'; -wa·n ⁂23 'complex plural suffix of nouns'; -wa·n ⁂62.1 'specific plural suffix of pronouns'; -wi ⁂311.2 'verb stem formative'; -wi ⁂324 'noun stem formative suffix'.

-ya ⁂316 'verbalizer suffix'; yeh ⁂133.1 'third person pronoun root'; -yo ⁂323 'noun stem formative suffix'; -yo· ⁂52 'locative suffix'.

ZERO- ⁂133.3 'third person verb subject prefix'; -ZERO ⁂211.1 'present tense suffix'.

-? ⁂212.1 'punctiliar past suffix'.

AN ACOUSTIC ANALYSIS OF SHAWNEE: II*

NANCY P. HICKERSON
INDIANA UNIVERSITY

2. Consonantic spectra
2.1. Single-quality consonantic segments
2.1.1–4. Plosives
2.1.5–8. Fricatives
2.1.9–12. Resonants
2.1.13. Glottalization

2. By consonantic we mean spectra in which all or part of the frequency range is empty of harmonic energy;[1] secondary criteria may be that harmonic energy is weaker at the frequencies at which it is present than in adjacent vocalic segments; that harmonics appear discontinuous with those of adjacent vocalic segments; that noise pattern is present, over all or part of the spectrum. Our primary segmentation, that of vocalic vs. consonantic spectra, is thus made on the basis of sudden breaks or interruption of vocalic spectrum. We have already suggested a secondary segmentation of vocalic spectrum, on the basis of rapid shifting of formant position. Consonant spectra will be similarly subdivided as (2.1) Single-quality, and (2.2) multiple quality segments.

2.1.1. The spectrum of *p* is complex, consisting of a relatively long portion which may, intervocalically, contain faint harmonics in the lower frequencies; a release band or spike and a very brief interval before the onset of a following segment.

Intervocalic *p* segments measured range in total duration time approximately between 7 and 17 centiseconds. When preceded by a vowel, the onset of a *p* segment is marked by the sudden cessation of most harmonics—usually all above 1200 or 1500 cycles. Any which continue are markedly reduced in power. This harmonic spectrum may be fairly clear or it may be quite faint; in any case, the range of frequency typically narrows, with only a trace or with silence (empty spectrum) just before the release band. The distribution of harmonic power, when it is present, varies slightly according to the preceding vowel. After *a*, power may continue in the range 400–1400 cycles, narrowing finally to 700–800 cycles, which is generally the darkest part of the spectrum. There may be some power higher, strongest at 1600 cycles. After *e*, power may be present initially from 300 to 1200 cycles, with the strongest and most enduring part of the spectrum about 500–600 cycles; there may be a light distribution of power at 1500–2000 cycles. After *y* or *i*, there may be power from the baseline to 1000 cycles, narrowing to 500 cycles; traces of upper frequency power have been observed, at about 2000 and 3000 cycles. In general, the location of the strongest part of this harmonic spectrum seems to be covariant with the first formant of a preceding vowel, but about 100 cycles higher than that formant (as influenced by *p*).[2]

The release spike of *p* is light and the interval between spike and vowel onset is typically brief. This interval ranges approximately from 1 to 3 centiseconds and has no apparent general correlation with total duration of the segment (though the short-

* For Part I see IJAL 24.20–29. [Editing of Part II of this paper was done by Fred W. Householder, Jr. CFV]

[1] Compare Jakobson, Halle and Fant, *op. cit.*, page 19; consonantal phonemes "are characterized by the presence of zeroes. . .".

[2] Auditorily, this harmonic part of the consonantic segment (in *p* and other obstruents) can be—with difficulty—perceived as prenasalization.

est segment measured is the most lenis and the longest has a relatively strong release). The following list gives total duration time and duration of spike from several *p* segments:

total duration time[3]	release interval[4]
7.0 cs.	zero
10.0	1.5, 1.5, 1.5, 3.0 cs.
10.5	2.0
12.0	2.0
12.5	1.5, 1.5
14.0	2.0
17.0	2.5

The explosion of the labial plosive is brief and may appear as a light but continuous distribution of energy from around 500 to 2000 cycles. The most consistent area of maximum energy lies between 1500 and 2000; this maximum is placed nearer 1800–2000 for *i*, slightly lower, nearer 1600–1800 for *e*, *a*, and *o*. There may be higher, less intense maxima at around 2500 and 3000 before *i*, 2200 and 2700 before *e*, and one at 2000 before *a*. Sporadically, before all vowels, a low maximum may occur between 300 and 500 cycles.

p affects the formants of the four simple vowels as follows:

i formants 2 and 3 are bent strongly downward; formant 1 is less obviously affected, and may appear level or be bent slightly down. All three formants of *e* are bent down, formant 2 strongly, formant 1 less so. The third seems to be bent only slightly. All three formants of *a* also show a definite downward bending. Formant 2 is strongly bent, but not so strongly as that of *e*. The same is true of *o*; the first formant is bent moderately, the second and third slightly.

Hypothetical F1 and F2 positions ('hubs') which would account for the bending of first and second formants of vowels are located near 400 and 900 cycles.

2.1.2. The duration time of *t*-segments measured ranges between 9 and 17 centiseconds; the average is about 13 cs. Harmonic power in post-vocalic *t* is distributed as follows. Following *a*, there is some power from about 500 up to 1800 cycles, continuously or in two concentrations. The lower harmonics narrow toward 8–900 cycles, which is also apt to be darkest part of the spectrum; another concentration may appear at 16–1700 cycles. After *o*, there may be harmonic power from the baseline to about 1000–1500 cycles, narrowing to about 600 cycles, where power is strongest. A second maximum may appear at 1500. After *e*, the strongest harmonic spectrum is about 650 cycles; after *i*, about 500. In these last two, distribution is less general and not so strong as after the other vowels.

The interval between release and vowel onset ranges between 1 and 3 cs., as with *p*; the most frequent duration is around 1 ½ cs. The following list gives total duration time and release interval of a number of segments:

total duration time	release interval
9.0 cs.	1.5 cs.
10.0	1.0, 1.5, 1.5, 2.0, 2.0
11.0	1.5
13.0	2.0
14.0	1.0, 1.5, 2.0
14.5	1.5
15.0	2.0
15.5	1.5
16.0	1.0, 2.0
17.0	2.0

The energy in the *t* release is distributed between around 800–1000 and 3000 cycles.

[3] When 'total duration time' of consonants is given, it includes the time between the terminus of one vowel to the onset of the next, including prenasalization, stop interval, explosion band, aspiration, silent interval, etc.

[4] Evidence of aspiration between explosion and vowel onset is almost always absent. The explosion band itself generally measures less than 2 centiseconds, the following interval about the same. Finer measurement has not been attempted because variation in spectrograph setting and measurement error might well account for whatever variation emerged.

The picture is very similar to the *p* explosion, with the most frequent area of maximum energy at around 2000 cycles—1800 to 2000 before *a*, about 1700 before *o*, 2100 before *e*, between 2000 and 2500 before *i*. Another maximum usually falls around 3000, and a third at about 1000 cycles. However, there is usually some indication of energy below the first and between the first and second maxima, so that the whole range from about 500 to 2000 cycles may be filled.

The following effects on vowel formants have been observed: in *i*-segments, the second formant is bent down slightly, the third seems about level; the first also does not generally seem to be bent. *e* formants all seem to be bent down slightly. The first formant of *a* is bent down, the second and third rather sharply up. The same is true of *o* formants, with the second being more strongly affected than in *a* segments. Hypothetical F1 and F2 positions would be near 400 and 1750 cycles.

2.1.3. *k* is quite variable in total duration and in variety and duration of release. Like *p* and *t*, *k*'s are lenis, with no sign of aspiration; but sometimes there may be two or even three separate explosion bands. Further, the explosion may occur sooner than that of *p* or *t*, and the open interval between explosion and vowel onset may be relatively long —as much as half the total duration of the segment.

As with *t*, a light vertical band occasionally appears shortly (e.g. 3–5 cs.) after the beginning of otherwise typical *k*-segments; these bands are similar in appearance to but lighter than the release bands.

Harmonic power in *k* segments is distributed very much as in *t*. After *a*, there is harmonic power from 500 to 1500 cycles; this is heaviest and endures longest at about 700 cycles, secondarily between 1200 and 1600 cycles. After *o*, the main concentration of power is between 500 and 600, sometimes around 1200 cycles. After the front vowels the usual occurrence of harmonic power, generally light, is around 500 cycles; traces in higher frequencies are sporadic.

The following list illustrates the variety of duration and release interval times.

total duration time	release interval
7.0 cs.	—, 4, 4.5 cs.
7.5	—, 2.5
8.0	—, 3
8.5	1, 2, 3.5
9.0	—, 2, 5, 5.5
9.5	3, 3.5
10.0	3, 3, 3, 3.5, 4, 4, 4
11.0	—, 2, 3, 3.5, 4, 5
11.5	6.5
12.0	—, 3.5
13.0	—, —, 3, 6
14.0	—, 4

While some power may be distributed over a wide area in the *k* release, there is usually one clear maximum with its placement apparently conditioned by the hub of the following vowel. Before *i* (and *y*) there is a maximum between approximately 2800 and 3300 cycles; there may be very slight aspiration here, filling the open interval between release and the beginning of the second formant of *i*. The explosion power is not highly concentrated at this maximum but is diffuse from around 500 cycles. There is, however, an open gap between 2000–2500 cycles and the maximum.

Before the low central vowel, *a*, there is a maximum of energy between about 1200 and 1600 cycles, and some distribution of energy just below, from, say 700 to 1200 cycles. There is occasionally a slight distribution of energy at 1000 cycle intervals above the maximum, at about 2100 and 3000 cycles (cf. *t*, above).

Before *e* the maximum seems slightly higher, between approximately 1400 and 1700 cycles, with a secondary area around 3000 cycles and lighter streaks sometimes appearing below 1400 and above 3000 cycles. Before *o* the maximum is lower, between 900 and 1300 cycles. Again, there may also

be a very slight general distribution of power both below and above the maximum, but no clear pattern or secondary maxima appear.

k affects neighboring vowels as follows: The first formant of *i* remains level or is bent slightly down; the second formant is bent slightly upward and the formant band often appears to merge with the third formant which remains level or is even bent down. All three formants of *e* are bent down slightly. Formant 1 of *a* is level or bends slightly down. The second formant is not consistently influenced in either direction, but the predominant tendency seems to be upward; the third formant is not strong and seems rather diffuse, but appears to parallel the second. The first formant of *o* is bent down or level, the second up slightly; again the third is difficult to trace, but it seems to parallel the second. Hypothetical formant positions would be about 400 and 3000 for k_1 (the type occurring before *i* and *y*), 400 and 1500 for k_2.

2.1.4. Plosives, as a class, are segments consisting of a relatively long interval in which power is either absent or is lightly distributed and diminishes with time, followed by a relatively short burst of concentrated power, the release or explosion. All plosives are lenis—indications of aspiration are rare and ephemeral. *p* and *t* are briefly and lightly exploded. *k* may have a more complex release, with two or more separate bands. The overall duration of plosives is highly variable, from approximately seven centiseconds to thrice this; for the most part (except when the release of *k* is prolonged) this variation is in the initial interval, not in the release, which usually has a very narrow range of variation.

Post-vocalically, harmonic power may be continuous from the vocalic segment through part of the duration of the plosive segment. This is strongest initially and decreases with time. The frequency distribution of harmonic power in plosive segments is generally below 1500 cycles, and is strongest approximately 100 cycles above the terminal position of the first formant of the immediately preceding vowel. After *a* and *o* a similar but weaker concentration may be seen at or near the second formant frequencies of the preceding vowel. (In other words, the distribution of power, while weaker than for an ordinary vowel, is approximately as for the preceding vowel as influenced by a nasal consonant.)

The explosions of *p* and *t* are characterized by broad distribution of energy, that of *k* by relative concentration of energy. The maximum of *p* is located between 1500 and 2000 cycles, varying slightly with the following vowel; a second maximum may appear at 500 cycles or lower. *t* follows a pattern of distribution of energy at 1000 cycle intervals—at about 1000, 2000 and 3000 cycles.[5] The most consistent maximum is at 2000. *k* has maximum energy at approximately 3000 cycles before *i*; before other vowels the maximum falls between approximate'y 1200 and 1600 cycles, the variation apparently conditioned by the following vowel.

Very rarely a single band, similar in appearance to a release band, occurs a few (3–5) centiseconds after the onset of (otherwise typical) *t* and *k*. In both cases, energy in this band is distributed between 100 and 3000 cycles; some *k* occurrences suggest a maximum at about 1300 cycles.

All three plosives have approximately the same influence on the first formant of adjacent vowels—F1 of *o*, *a* and *e* are bent down, while that of *i* remains level or may also be bent down (a low first formant is difficult to trace precisely). They have in common, then, a F1 at about 400 cycles. Second formant behavior varies; the F2 hub of *p* is relatively low, and all second form-

[5] A tendency frequently noted, c.f. G. Fant, "Transmission Properties of the Vocal Tract," part II, *Acoustic Laboratory*, M.I.T. Quarterly Progress Reports, Oct.–Dec. 1950, p. 17; Fisher-Jörgensen, *op. cit.*, p. 49.

ants are bent down by *p*. It is located, therefore, around 900 cycles (at the highest).[6] The *t* F2 hub is above the level of the *a* nucleus, level with to slightly below *e*; it may be estimated at 1750 cycles. *k* has one variant (k_1) with F2 hub around 3000 cycles, as the *i* second formant is bent up. Before other vowels, it (k_2) has a F2 hub higher than the *o* nucleus, and apparently somewhere between *a* and *e*—perhaps relatively closer to *a*, as *t* is to *e*. We will estimate it at 1500 cycles.

2.1.5. The range of duration of *s* segments is between approximately 10 and 17 centiseconds; most fall in the vicinity of 12–13 cs. (The actual noise spectrum is usually preceded and followed by a silent interval of 1 or 2 centiseconds; these intervals are included in the measurements given.) A small sample of *s* segments has the following distribution of duration time measurements:

duration time	number of instances
10 cs.	1
12	5
12.5	2
13	4
14	3
16	3
17	2

The average duration time is 13.5 centiseconds, about 3 cs. longer than the average for stops (see above).

The usual *s* spectrum includes a dark band about 300 cycles (or more) wide, in the vicinity of 2000–2500 cycles, and a cloudy area above 2800 cycles. Two bands, one around 2800–3000 cycles, the other around 3500 cycles, usually stand out in this area, which continues up to 8000 cycles, with diminishing intensity above 6000 cycles. There is at times some power between the lower band and the cloudy area, but always weak. We will refer to the lower band and the two which stand out in the large cloudy upper area as formants. The *s* formants are, if projected through the hiatus between the vowel and *s* segments, continuous with the second, third and fourth formants of preceding and following vowels.

In post-vocalic *s* segments, some harmonic power may be present in lower frequencies, co-occurring with the upper-frequency fricative spectrum. The distribution of this power is as in the plosive segments, perhaps more similar to that in *k* than to *t*—after *a*, the strongest concentration of power is approximately 700–800 cycles, and occasionally around 1400 cycles. After *o*, *e*, and *i*, the most consistent occurrence is in the vicinity of 500 cycles. Harmonic spectrum seldom endures for more than half the duration of *s*; as with plosives, it seems strongest after *a*, weakest after front vowels.

The high front vowel approaches *s* with a first formant approximately level, a second formant level or falling slightly, third formant level. The lower front vowel, *e*, has F2 bent upward by *s*. F1 is level or bent slightly down. Comparison of the effects of *s* on *i* and *e* segments is sufficient to localize its F2 hub in the vicinity of 2200 cycles, being slightly below the *i* and considerably above the *e* hub. The F1 position is, as for the plosives, at or below the *i* level, about 400 cycles.

a, then, has the second formant bent sharply up; formant 1 is bent down. The second formant of *o* is bent up drastically by preceding or following *s*; formant 1 is again bent downward.

The formants in the *s* spectrum itself are located at frequencies obviously related to those of formants of the preceding and following vowels. Between two high front vocalic segments, the first *s* formant is horizontal, at 2500 cycles or slightly below, connecting the second formants of the vowels; the second is located at 3100 or 3200 cycles, connecting vowel third formants; etc. Preceded and followed by *a*, the first *s* formant is again level, at about 2000 cycles,

[6] c.f. Danish /p/ at 1300 c.p.s., Fisher-Jörgensen, *op cit.*, page 52.

the *a* second formants rising to this point; the second at about 2900, the third at 3300 (when the second and third can be separated). Between *a* and *i*, the first *s* formant rises, from around 1700 to 2200 cycles, connecting the *a* and *i* second formants (at approximately 1700 and 2400 cycles, respectively); the second and third *s* formants appear inseparable, above 2800 cycles.

In brief, the lowest *s* formant more or less directly connects the second formants of the adjacent vowels, except when both second formants are below the hub frequency of the *s* (e.g., when both are *a*). In the latter case, the two vowel formants are pulled up, and the resonance band is located at approximately the *s* F2 hub (slightly above 2000 cycles). The upper *s* formants, when separable from their noisy matrix, connect vowel formants three and four (both are usually between approximately 3000 and 4000 cycles).

2.1.6. *θ* segments range in duration from around eight to sixteen centiseconds. Most segments fall around 11 or 12 centiseconds, however. The distribution of duration in a small sample of *θ* segments is as follows:

duration time	number of segments
7 cs.	1
8	2
9	1
10	2
11	4
12	4
13	3
14	0
15	2
16	1

The friction spectrum of *θ* is usually weak; sometimes friction seems almost absent, so that the most striking difference between *θ* and a dental plosive is the absence of an explosion band. Usually, however, there is some distribution of friction in the lower frequencies, between zero and 2500 cycles. Another irregular concentration of energy, a cloudy blur of variable size, is present in the vicinity of 5000 cycles (from 4000 to 6000 when relatively heavy). Lighter traces between these two areas are sporadic—a light band at times appears at about 3000 cycles.

Harmonic power, too, is generally lightly present in post-vocalic *θ*; the distribution of this power appears quite similar to that in *s* and the plosives. Unlike *s*, which may have harmonic power in lower frequencies, friction in upper, *θ* often has both harmonic and frictional spectrum overlapping in the lower frequencies.[7] After *a*, there is some harmonic power at about 600–800 cycles, with friction above and below. Harmonic power seems to be strongest generally after *o*; harmonics up to about the fifth may be continuous from the preceding to the following vowel. The darkest areas are at about 500–700 and 1300 cycles. After *e*, the picture is similar, with concentrations at about 600–700 and 1800 cycles. After *i* the distribution is light, strongest at 450, occasionally present around 2600 cycles. In the last two cases, frictional power appears between the lower and upper concentrations of harmonic power.

i, preceding or following *θ*, is influenced in that the second formant is bent somewhat down (apparently slightly more strongly down than by *s*). The third is also bent slightly down; the first is level or bent up a little. *e* has the first, second and third formants bent slightly down. *a* formant 1 is bent down, while the second is about level. *o* formants 2 and 3 are bent quite definitely up; the first is level or bent down slightly. The F2 hub, then, is slightly below the *e*, about level with the *a* nucleus. It can be estimated at about 14–1500 cycles. The F1 point is approximately that of the plosives and *s*, perhaps slightly higher since the *o* first formant remains level.

The primary focus of fricative energy in the *θ*-spectrum itself seems to be the vicinity of the first and second formants of the preceding and following vowels. When

[7] When light and scattered the two kinds of energy may be indistinguishable in spectrum.

preceded and followed by *i*, there is a general distribution of power between approximately 500—slightly above the level of the first formant of *i*—and approximately 3000 cycles. Otherwise the spectrum is a blank, but for a very slight blur continuous with the fourth formant of the preceding *i*. Between two *o*'s, the concentration of power is between 600 and 1700 cycles—from the level of formant one to just above the second formant—with nothing above or below except for slight traces at 3000 and 5000 cycles. When preceded by *o* and followed by *i*, the upper edge of the dark area rises from onset to terminus from 1500 cycles, continuous with the second formant of *o*, to around 2000 cycles, nearing the point of onset of the *i* formant. Again, there may be slight energy at 3000 and 5000 cycles.

In summary, the main concentration of energy in *h*—at times the only noise spectrum present—is in the area between the first and second formants of the adjacent vowels (and, perhaps, below the hub of a preceding consonant). Secondary foci of energy center around 5000 and 3000 cycles (the latter approximately continuous with vowel formant three).

2.1.7. The duration range of intervocalic *h*-segments is approximately six to sixteen centiseconds; in a sample of twenty segments the distribution is as follows:

duration time	number of instances
6 cs.	1
7	6
8	2
9	0
10	2
11	2
12	4
13	1
14	0
15	1
16	1

Energy in the *h* spectrum is usually weak, approximately as in the usual *h* and weaker than the *s*. There is usually—but not always—some frictional power above 500 and below 2000 cycles, and some at approximately 3000 cycles. The latter is predominant in a front vowel environment, while the lower frequencies are stronger in the vicinity of *e*, *a*, and *o*. Between two relatively low vowels, e.g. (aha), the power is concentrated between 500 and 1800 cycles, with a trace at 3200–3500. Vowel formants 1 and 2 are continuous with the lower power concentration; a darker band may stand out through this matrix to which the *a*-formants are more directly connected, between approximately 1000 and 1400 cycles. Between two high front vowels, (ihi), the noise spectrum centers around formants two and three of the vowels, between approximately 2500 and 3200 cycles. Two bands, continuous with the vowel formants, may be traceable through the *h* spectrum. Between unlike vowels, a band connecting vowel second formants, at least, can usually be traced. In the sequence (ohi) for example, one band directly connects the *o* first formant, which terminates at approximately 1000 cycles, with that of *i*, which has its onset at around 700 cycles (both vowel formants are thus bent upward); the power within the *h* segment bridges the gap between 1000 and 700 cycles. Another wispy resonance band continues the *o* second formant up to around 2000 cycles at a point approximately half way in duration of the *h* segment; the second half of the segment contains power continuous with formants two and three of the following *i*.

Very brief *h* segments, approximately 7 cs. in duration, in a front vowel environment, often contain harmonic power in the lower frequencies continuous with those of adjoining vowels. The *h* spectrum is thus primarily a noise spectrum, but in a minority of cases includes harmonic elements.

h has the following effect on the formants of adjacent vowels: The *i* first formant is usually unaffected or may even be bent down; at times it appears to be bent up when the other *h*-adjacent vowel is a low central

or back one (*a* or *o*). In other words, the effect of the *h* on the *i* first formant is essentially neutral; other factors may condition a shift in the formant position. Formants 2 and 3 are not bent by *h*.

The first formants of *e*, *a* and *o* are bent up rather sharply by *h*; second and third formants are generally not affected, usually measuring about median for the vowel in question.

Thus, in general, the main effect of *h* on neighboring vowels involves first formants, which are bent up in three out of four cases. For *i*, this seems to be usually absent. It may be that the first formant of *i* is more often bent down, but in some cases this can be reversed by other environmental factors.

2.1.8. Fricatives, as a class, are segments delimited by weakening and gradual cessation of harmonics (like stops) and characterized by the presence of a noise spectrum. By NOISE is meant power lacking a harmonic structure, distributed unevenly in cloudy, mottled or, when intense, irregular solid masses. At times, however, the frictional noise is either absent or so weak as to produce no spectrographic record;[8] the absence of a release or explosion band, characteristic of plosives, can validly be used as a secondary criterion for the identification of fricatives.

The approximate range of variation in duration for fricative segments measured is six to sixteen centiseconds. Only *h* segments are as brief as six cs., however; the lower limit for *s* seems to be around ten cs. The maximum, on the other hand, is approximated by all three fricatives; it does not approach the maximum found for plosives, 21 cs. The duration range of variation, then, is more limited than that of stops, and of fricatives *s* has the narrowest range.

Of the three *s* is the most powerful and the power is present at higher frequencies than is characteristic of the other two. Both *θ* and *h* are typically relatively weak, though variable; *θ* presents at times as strong a spectrum as *s*, and, too, some *h* segments are considerably stronger than the typical *θ*.

Each of the fricatives has a characteristic distribution of frictional power within its spectrum and a characteristic effect on formants of adjacent vowels. The areas of maximal energy in *s* segments are located between 2000 and 3000 cycles, generally with traceable resonance bands continuous with the second, third and fourth formants of adjacent vowels. Of these three bands, the upper two are often the stronger, and additional power in the spectrum lies in the upper frequencies, extending to around 6000 cycles. The center of gravity, then, is quite high—the friction is absent from the spectrum below 2000 cycles. In *θ* segments, the area of maximal energy lies between approximately 500 and 2500 cycles, above the level of the first and below the second formant of adjacent vowels; a secondary maximum may appear at 5000 cycles (when total power is relatively strong). In general, resonance bands do not stand out in the *θ* spectrum, though the general configuration of the noise area may appear to reflect environmental influences. The *h* spectrum is usually concentrated below 3500 cycles. There are two distinctive patterns of power concentration. When *a*, *o*, or *e* is adjacent, energy in the fricative spectrum is concentrated in the lower frequencies, above 500 and below 2000 cycles; a secondary concentration may occur in the vicinity of 3000. Adjacent to *i*, the primary concentration is above 2500 cycles, continuous with formants two and three of *i*. Resonance bars are often traceable in *h* spectra, joining the vowel formants.

Some harmonic power is generally present initially in post-vocalic *s* and *θ*, continuous with formant 1 and at times formant 2 of the preceding vowel. In relatively brief *θ* and *h*, lower harmonics are at times con-

[8] Sometimes friction profile may show up in the amplitude display but be imperceptible in the spectrum.

tinuous between preceding and following vowels.

The *s* F2 hub is somewhat higher than that of *t*, located at approximately 2200 cycles, or higher than formant 2 of all vowels but *i*. *θ* has a F2 hub below that of *t* and near the k_2 hub, at around 1500 cycles. We cannot localize an *h* F2 hub on the criterion of vowel second formant behavior, as *h* does not generally appear to affect second formants. However, one *h* spectral pattern—h_2—has the definite effect of greatly raising *e*, *a*, and *o* first formants; a first formant hub could be located above 1100 cycles. The other pattern, h_1, is associated with *i*; either there is no typical effect on *i* formants or else the first formant is lowered.

2.1.9. Almost all *m* segments fall within a narrow range of between seven and ten centiseconds in duration. Duration times in a sample of thirty segments are as follows:

duration time	number of segments
6 cs.	1
7	5
8	9
9	9
10	5
15	1

A formant pattern differing from that of adjacent vowels, and the fading out of upper and, at times, intermediate, harmonics, clearly demarcate *m* segments. The basic *m* spectrum consists of two formants, the lower always falling below 600 cycles, the second above this frequency and below approximately 2000 cycles. There may be some power distributed in frequencies just above and/or just below these formants, and occasionally a light distribution of power is seen at approximately 2500–3000 cycles.

m has the universal effect of bending vowel second and third formants down and first formants up. *m* thus has approximately the same F2 hub as *p*, around 1000 cycles or lower. Its effect on first formants, however, is the reverse of that of *p*; the F1 point must also be around 1000 cycles.

The placement of *m* formants seems to vary according to adjacent vowels. This effect is more obvious on the upper formant, less so on the lower. Preceded and followed by a high front vowel, the first formant of *m* usually falls around 400 cycles (±100, or with power distributed generally between 300–500), the second around 1500–1600; power may be present at 2500. Preceded and followed by *a*, the first formant is located at about 400 cycles; the second at 1100—considerably lower than with adjacent *i*. The power continues, diminishing, up to about 1400, however. *o* seems to bring the second formant even lower, to 900 or 1000 cycles.

Between unlike vowels, the second formant of *m* reflects, in a rather subdued way, the qualities of the two vowels. In the sequence (ema), for example, it falls from approximately 1300 to 1000 cycles. In (ame) it rises from 1200 to 1600.

The formant bars in the *m* spectrum are uniformly lower than the corresponding formants in adjacent vowels; it is the sudden downward shift in formants which clearly demarcates *m* segments. However, both vowels and nasals are harmonic segments, and the transitions between them, though sudden, are traceable. Thus, the first and second formants of *i* and *e* join formants 1 and 2 of *m*, and the third is traceable to the vicinity of 2500 cycles where a formant seems to be potential, though represented only vaguely or not at all on spectrograms. Of *o* and *a*, the first formant joins the first of *m*, while the second and third both appear to be drawn toward the vicinity of 1000–1200 cycles, where two formants of *m* may come together (though appearing as a single band in spectrograms).

2.1.10. *n* is very similar to *m* in duration, spectrum, and in some of its effects on vowel spectra. Like *m*, the majority of *n* segments are about eight centiseconds in duration,

the general range being between seven and twelve cs.

duration time	number of segments
7 cs.	2
8	9
9	5
10	2
11	1
12	1

Like *m*, *n* can be said typically to have a completely harmonic spectrum containing three formants, the uppermost of which is often weak to the point of invisibility in spectrograms. The second is generally weaker than the second formant of *m*; only the first is always clearly traceable.

The lower formant of *n* is almost always located below 500 cycles; 450 is perhaps an average. Power in the lower frequencies is usually distributed over three harmonics between approximately 300 and 500 cycles. The second formant ranges between 1600 and 2200 cycles, centering around 1800. This range is not unlike that of *m*, but the *n* formant never goes as low as that of *m* may in certain environments. Traces of a third formant may appear around 3000 cycles. In a front vowel environment, the second formant lies in the vicinity of 2000 cycles or slightly above—up to 2200 with *i*, around 1800 with *e*; formant 1 centers around 400–450 cycles. *a* and *o* lower formant two to around 1600 and formant one may be raised by 100 cycles, to 550. As with *m*, the second formant may rise or fall, according to vowel environment. In (ena) it falls from 1850 to 1600 cycles, for example; in (oni) it rises from 1600 to 2000.

n raises first formants of all adjacent vowels. As with *m*, the vowel formants are located higher than those of the nasal, and wispy, almost vertical connecting lines between the two are traceable. *i* has a very high first formant, while the second is bent down rather strongly. *e* first formant is raised and the second is about level or bent down slightly, to about 1900 cycles, the level of the *n* second formant. *a* and *o* similarly have high first formants and both have the second formant bent up to or toward 1600 cycles. The *n* F2 hub is approximately that of *t*, slightly below the center of the *e* area, in the vicinity of 1800 cycles. However, while *t* generally lowers, *n* raises first formants, like *m*, toward about 1000 cycles.

2.1.11. Like *m* and *n*, *l* segments are relatively brief in duration, with little variation. The range is between five and ten centiseconds; most segments fall around 7–8. The chart shows the distribution of *l* segments in the duration range and also indicates the occurrence of a brief initial stop interval (see below).

total duration time	initial stop interval	number of segments
5 cs.	—	1
7	—	5 } 7
7	1 cs.	2 } 7
8	—	4 } 9
8	1	5 } 9
9	—	2
10	—	1

The general *l* spectrum is similar to that of *n*.[9] The first formant is located at around 400–500 cycles, more frequently 500. The second formant is located at approximately 1400–2000 cycles. The area between formants 1 and 2 often registers some power and, though the first usually is clear, the specific location of the second formant may be at times a matter of hypothesis. The frequencies above the second formant are usually blank, at least up to around 2800 cycles. A third formant may appear between 2800 and 3200 cycles.

l segments tend to increase in spectral intensity from beginning to termination, and some clearly have an initial silent interval of about one cs. duration. This is sometimes

[9] /n/ and /l/ are the phonemes most frequently confused auditorily.

combined with an explosion or release band (generally at approximately 2000 cycles). In other words, the *l* harmonic spectrum may be preceded by a very brief voiceless stop, which may or may not be exploded. When exploded, the maximum is level with the *l* second formant. This stop is brief, compared to the following continuant, and is hardly perceptible to the ear.

Both the explosion noise of the stopped part of the *l*, when present, and the position of the second formant seem to vary with the quality of adjacent vowels, being relatively low or relatively high accordingly. After *a*, there may be a light initial explosion at about 2000 cycles. Between two *a* segments, e.g. (ala), there is often a solid light distribution of power from the first *l* formant, at 450 cycles, to 1400–1600 cycles; formant 2 may or may not stand out at the top of this area. The third formant is located at around 3000 cycles. Adjacent to *o*, the first formant is between 400 and 500 cycles, the second at 1200–1400 (lower than with *a*). After *e* no stop interval has been observed, but there may be an initial blurring of the second formant, which is located at around 1600 cycles. After *i*, there is a stop interval and an explosion of 1–2 cs. duration, centering around 2000 cycles (below 2600). The first formant is around 500 cycles, the second around 2000, continuous with the lower edge of the explosion noise. The third is again around 3000 cycles.

Between *e* and *a*, (ela), the second formant of *l* falls from around 1700 to 1400 cycles, connecting the vowel formants. In (ilo), it falls from around 1400 to 1200. In (ale), the second formant rises from 1500 to around 1800 cycles.

The formants of adjacent vowels are affected by *l* as follows: the first formant of *i* is level, continuous with the formant in the *l* segment, or bent down very slightly; the second is bent down. *e* formant 1 is bent down, and the F2 is generally bent down slightly. F1 of *a* is bent down, F2 often seems level, sometimes is bent slightly up. *o* formant one bends down, formant two up. The *l* F2 hub is obviously very close to that of *t* and *n*, but apparently slightly lower, closer to the *a* territory, perhaps between 1600 and 1700 cycles. The F1 hub is near or slightly below the *i* nucleus, similar to that of the plosives.

2.1.12. We group as resonants spectra which are harmonic, but which are clearly delimited as segments by the sudden disappearance of harmonics at certain frequencies (by contrast to adjacent vowels) and a general diminution of power at all frequencies. Additional criteria are a general downward shift of formants in *n* and *m* and, at times, a brief initial stop interval in the case of *l*.

The three resonants are alike in having a duration range generally restricted to between seven and ten centiseconds, the average being around eight. This is both a narrower range of variation and a lower average duration than that of any other group of consonants.

Resonants have a formant pattern not unlike that of vowels; at least two formants are usually traceable. However, power is in general reduced in these segments, and, too, formants do not appear to be as concentrated as in vowels. In other words, formants in resonant segments are weaker and more diffuse than vowel formants. In general, first formants fall at around 500 cycles or below, second formants between 1000 and 2000 cycles. Third formants are generally weak or absent. *m* and *n* have in common a very diffuse first formant which probably centers around 400 cycles, apparently somewhat lower than the usual *l* formant. *n* and *l* have in common a second formant in the approximate range 1500–2000 cycles—the *n* range may average 100 cycles higher than the *l*—by contrast to *m* which has a low second formant.

m and *n* have the common effect of strongly raising the first formant of adjacent vowels, to a F1 position around 1000 cycles, while *l* has an approximately neutral

or slightly lowering effect. *m* and *n* also have the effect of diffusing the power in first formants of adjacent vowels (nasalization). *m* lowers all second formants, while *n* and *l* raise those of *a* and *o*. Thus, *m* has a low F2 hub, approximately 900 cycles, while the *l* and *n* F2 hubs are both in the vicinity of 1800 cycles (the *l* perhaps slightly lower). *m* and *n*, then, share first formant effects and contrast in their effect on second formants of vowels; *l* and *n* share second formant effects and contrast in their effects on first formants of vowels.

2.1.13. One may better speak of 'glottalization' than of 'glottal plosive' or 'glottal stop', terms which suggest a single spectrum segment comparable to other consonantic units. The spectra described here are harmonic, consisting of vocalic and resonant spectra, as or similar to those described above, **1.2**, **1.3**, **2.1.9–12**, which are vertically striated by brief bands of silence. That is, though the glottal effect may be considered in fact a series of stop (silent) intervals, these are recurrent and so brief that the harmonics and formants remain clear and traceable. We will describe glottalization as an additive feature, superimposed on spectra which fall within limits which we have already defined, or sequences of these.[10]

Glottal bands—extremely brief intervals with a duration range of less than one centisecond to around three centiseconds (at the most)—intersect harmonic spectra and divide them into pulses of a similar duration. Thus, a ten centisecond stretch of glottalized spectrum may contain from about three to six such pulses.[11] Another variable, characteristic of these segments, is a general fading of harmonics peripheral to the formant frequencies. When present, this seems to be a terminal feature, the harmonics fading toward the end of the glottalized segment. Formants may be concurrently visibly weakening, especially those of resonant consonants. Glottalization, then, combines repeated brief interruption of harmonic spectrum with a progressive attenuation of the energy in the segments affected.

If we compare the formant patterns of segments affected by glottalization with those of unglottalized segments in order to see whether ʔ has qualitative effects on spectra with which it occurs, we find that, in every case, the paths of glottalized segments fall within the vowel domains already established; however, glottalization appears to have a definite effect on segments in some of the domains. The qualitative effect on *a* and *o* segments is clear—termini in these domains have first formants located approximately one to two hundred cycles higher than in equivalent non-glottalized occurrences. Thus, for example, in (moʔθ) the vowel formants begin at 500 × 700 cycles, about as in a non-glottalized vowel following *m*, but terminate at 750 × 1300—the second formant position about that of a non-glottalized *o* preceding *θ*, the first about 100 cycles higher.

Glottalized front vowels are infrequent in occurrence and qualitative effects of glottalization on these are less clear than for *a* and *o*. It may be that the striation of the harmonic spectrum is the only glottal effect on these segments. Glottalization effects on resonant spectra consist of vertical striation, as with vowels, and a considerable weakening of the (already attenuated) power concentration in such a way that the formants appear lighter while their distribution over a given range of frequencies is not restricted.

[10] By contrast to *h* which, while having a rather similar effects on formants of adjacent vocalic segments, interrupts the harmonics completely.

[11] Some of this variability may be apparent only, for the reasons stated in footnote 4, above.

NOTES AND REVIEWS

ESSAYS IN LINGUISTICS. By Joseph H. Greenberg. Chicago, The University of Chicago Press, 1957. vii + 108 pp.

W. P. LEHMANN

In his eight essays on linguistics Greenberg deals with two approaches to language, with language 'as a system of signals conforming to the rules which constitute its grammar or as a set of culturally transmitted behavior patterns shared by a group of individuals' (1 [Page references to Essays in Linguistics will be enclosed in parentheses]). The first approach Greenberg undertakes to illuminate by discussing characteristics of sign systems; the second, by discussing two problems in historical linguistics and three in ethnolinguistics followed by one in general linguistics. These topics, by the knowledge they require in various branches of linguistics and in related fields, such as branches of mathematics and anthropology, place considerable burdens on the author, but Greenberg has apparently shouldered them in the belief that 'further substantial progress in linguistics requires the abandonment of its traditional isolationism (v). His success depends on his mastery of these fields and his skill in presenting their salient features.

Greenberg's first essay, Language as a Sign System, with three short appendices and a glossary of technical terms is an exposition of the theory of sign systems, starting from Morris' essay of 1935 and making use of set theory. Procedures relevant for linguistic analysis are then discussed in closely reasoned argument interspersed with mathematical formulae. These procedures are applied in the second chapter, The Definition of Linguistic Units, to the morpheme and the word. Though he mentions phonemes Greenberg does not identify morphemes or 'words' by them; he himself calls 'the absence of any treatment of phonemic theory [and of any over-all discussion of meaning]' the 'most serious omission' of his book (v). To judge by his reference to the Eskimo phonemic system as consisting of 18 members (33), Greenberg does not consider suprasegmentals phonemes. The morphs he deals with are only those composed of segmental phonemes. And the goal of his morphemic analysis is 'to determine how many morphs there are in each word' (27). The word accordingly occupies a central position in Greenberg's morphological analysis, and he devotes considerable space to defining words. In his definition he takes 'the existing usage as a point of departure and one to which our results must, in general, conform' (28). Tentatively defining words as 'non-interruptible units' and 'stretches between word boundaries' he uses morphs and a further intermediate unit, the nucleus, to determine word boundaries. 'A nucleus boundary is an infra-word boundary if there is a maximum to the number of nuclei that can be inserted.' 'It is a boundary between words . . . when there is no maximum to the number of insertions of nuclei, if there is "infinite" insertion' (30). Thus *the* and *farmer* are words because one can insert *cruel, ruthless*, etc. On the other hand, Latin *dominúsque* in *sérvus dominúsque venḗrunt*, though consisting of three nuclei, is one word. Greenberg considers his definition preferable to previous attempts because it 'resolves the contradiction between phonological and grammatical definitions,' adding that definition of words by 'stress or some other marker . . . sometimes leads to strange results.' The 'strange result' he cites is the analysis of Polish *Améryka* into a 'word' *Améry* + *ka*, if words in Polish were to be defined by penultimate stress. Presumably it is not a strange result to oppose 'existing usage' and consider 'words' followed by enclitics, like *dominúsque*, words. Forms like English

can't, couldn't, didn't in *he . . . come* must then be words, and possibly *am* is not, in the sequence *Am I late?* Nor would Japanese post-positions like *e, ga, wo* be words. Greenberg's definition then fails to meet the test by which he rejects Bloomfield's: 'it leads to some results not at all in agreement with the traditional notion' (28). Bloomfield however did not make the word his central morphological unit. It is unfortunate that anthropologists, to whom Greenberg has principally directed his essays, should be led to believe a quarter century after Bloomfield's Language was published that it is useful to make a graphemic unit into one's basic linguistic unit, whether or not one supports the attempt with argumentation based on set theory.

The first essay in historical linguistics deals with Genetic Relationships among Languages, the second with the problem of Linguistic Subgroupings. Greenberg seeks to analyze the methods used in formulating hypotheses of linguistic relationships in order to resolve conflicting classifications and to enable the non-specialist to evaluate alternative classifications. He says that the 'method for discovering valid relationships' rests on 'two main principles—the relevancy of form-meaning resemblances in morphemes to the exclusion of those based on form only and meaning only and the technique of group comparison of languages' (40). The first principle is obvious; by the second Greenberg states that the earliest possible form of two languages be compared, e.g. Germanic with Indic, not English with Hindustani.

Greenberg's principles are good, as far as they go. His essays on historical linguistics would have profited greatly by reference to A. Meillet's La Méthode comparative en linguistique historique (Oslo, 1935). Meillet does not at all consider the 'compilation of comparative fundamental vocabularies' the 'swiftest and surest method' of establishing historical relationships. Even Grimm and Bopp did not. It was their great achievement to launch the principles discussed in Meillet's book. In comparing languages it is the systematic relationships and differences which are important. As Meillet points out there is no phonetic agreement between Armenian erku *two* and Latin duo, Skt. dvā, etc. But when one determines the systematic points of agreement and lack of agreement between Armenian and other Indo-European languages, the changes which took place between the Proto-Indo-European word for *two* and Armenian erku become apparent, and one can equate readily the Armenian and Latin words for *two*. If one simply listed the forms side by side, as is done in Table 2 (41) for a variety of languages, one would scarcely find a connection between the Armenian and the Latin forms. Since linguists working on languages with few or no materials from the past are devoting considerable time to the establishment of genetic relationships, it may be well to bluntly point out essential principles in historical linguistics. It is saddening to see comparisons of vocabulary of the sort that flourished before Grimm, with no attempt to establish systematic sound and form relationships. Meillet's short book would help other linguists arrive at better method than does Greenberg's.

Greenberg rightly points out the importance of solving problems of linguistic subgroupings, e.g. the subgrouping of Germanic. His discussion deals chiefly with the principles to be used, principles which he states were presented by Brugmann in 1883. Since opinions differ on the problem of subgrouping even in Germanic and other languages where there is no shortage of materials, it requires some optimism to expect results in less well documented areas that will be of value to ethnologists and culture historians. It is well to stress with Greenberg the importance of the problem and the difficulty of solution. It would also be comforting to see an essay on the principles help clarify a specific problem in subgrouping.

The last essay, Order of Affixing: A Study in General Linguistics, attempts to account for a statement of Sapir's that 'prefixing is far less frequent than suffixing in the languages of the world' (89). In his discussion making use primarily of inflectional affixes, Greenberg adduces psychological factors which favor choice of suffixes, and tentatively suggests typological sequences for the development of languages.

The three essays on problems in ethnolinguistics concern themselves primarily with the definition of terms. The essay on Language and Evolutionary Theory examines how the concept of evolution has been applied in linguistic study and comes to the conclusion 'that it is not language as such which evolves but rather communication in general' (65). The essay on Language, Diffusion, and Migration examines the various types of linguistic classification and the interaction of languages, with a view to clarifying methodology dealing with similar situations in other cultural study. The essay on Structure and Function in Language deals primarily with the definition of function and its use in mathematics, biology, and the social sciences, and branches of linguistics.

The problems raised in these three essays are obviously central to linguistic study. One would welcome a close examination and definition of them. But in definition precision and clarity are essential. These qualities are lacking in Greenberg's writing. Possibly one could not expect precision from a text which includes the following statements: 'Actually, much ambiguity is quite harmless and even a useful conservation of vocabulary resources. For example, the use of the term "case" both in grammatical discourse and in the law courts will presumably mislead no one' (63). In reading these essays we find 'case' used an average of once a page, culminating with eleven occurrences on page 88. Conservation indeed, but one's key words should surely be unambiguous. As an instance of lack of precision and clarity we might cite the abundant use of 'this' with unclear antecedent, e.g.: 'From the phonemes which compose the units we cannot predict their meaning. This is the well-known principle of the arbitrariness of the relation between sound and meaning. In this case we have seen, in the discussion of linguistic relationship. that, although prediction is not possible, we can in certain cases involving sound symbolism arrive at better than chance results. This is the extreme instance of arbitrariness' (84). One can generally determine what Greenberg is attempting to say, but one receives little help from his method of saying it. And if one cannot rely on precision in the presentation of techniques which are adduced to assist in the 'abandonment of isolationism in linguistics,' one is better advised to spend time on existing treatises on linguistic method.

Since writings on linguistics do not abound in stylistic niceties, we may seem to be condemning one book for the faults of many of its fellows. But while the stylistic difficulties cited may be those most frequent, other characteristics scarcely contribute to recommending this book as the explicator of linguistic theory for the anthropologist. Few forms are cited, in contrast to works by Bloomfield, Jespersen, Sapir; and many of the cited forms are puzzling. In Table 3 (43), a list of nine words from languages of Europe, the German word for *one* is given as ajns, for 'two' as tsvai; the Russian word for 'one' as adjín, for 'three' as tr^{j}i; presumably we are to conclude that the differing transcriptions for the German diphthong refer to differing pronunciations, and similarly the differing transcriptions for Russian palatalization. The table may be a classic for its errors. The French given for *eye* is œl/jõ; elsewhere plurals are not given, except for the Estonian for *tooth* which is cited in the acc. pl. hambaid. The Irish word for *one* is given in the conventional orthography, aon; the word for *two* is given as dau, which one can explain only on the assumption that since words for *two* were not given in Buck's

Dictionary of Selected Synonyms Greenberg went to the sections of Brugmann's Grundriss which Buck referred to and copied down Brugmann's representation of the Old Irish form for *two*. The Estonian word given for *ear* means *ear of grain*, not the body organ. A full review of Table 3 would require considerable space. And one of the few sets of words cited is: 'Sanskrit *veda*, Greek *woida*, Gothic *wita*, all meaning "I know" ' (45); the first of these forms is correct, if one discounts omission of the accent.

References to earlier work in historical linguistics also fail to raise one's estimation of the book. Few Indo-Europeanists would even suggest that 'Proto-Indo-European might have had two futures with different semantic functions' (52). The locution 'changes stated by Grim in his first law' (51) is strange, as is a statement like the following 'The Latin form *quattuor*, "four," is a first-rate etymology because it is long, exhibits recurrent correspondences in most of its parts, occurs in every branch of Indo-European, and is semantically straightforward' (45). Again one can with good will determine what Greenberg means; but he is using an infrequent meaning of etymology, and quattuor can scarcely be said to occur in any Indo-European language other than Latin. Such inaccuracies disturb the course of the text, as do gratuitous remarks like 'poetry, which flourishes on vagueness' (64).

It is really regrettable that these essays have so many shortcomings, for their topics are important and they reflect a great deal of work on the part of their author. That much work was necessary before they were published can scarcely excuse their faults or recommend them to the general student who wishes to learn the techniques of linguistics. The virtues of the book may not have received adequate recognition here, nor have all its infelicities; but it is wearisome to battle through inexactitudes and errors to an author's contribution. We no longer have such a paucity of books in linguistics that haste in publication at almost any cost is necessary for progress.

We all look to progress in linguistics, but the essential for it is not dipping into other sciences for their techniques. Mastery of method in other sciences may stimulate progress, but the essential condition for it is that stated by Meillet for historical linguistics, 'il importe de préciser, de systématiser et d'étendre les recherches' (La Methode, 109). The works of the great linguists of the past, Paul, Meillet, Sapir, Trubetzkoy, Bloomfield do not support Greenfield's statement that linguistics has been traditionally isolationist. Though some of their facts and methods may be obsolete, their works on linguistic method and their specific studies still must be thoroughly exploited before any pretence can be made of further contribution to linguistic study.

UNIVERSITY OF TEXAS

CONCERNING THE VALIDITY OF LEXICOSTATISTICS[1]

JOHN A. REA

Glottochronology,[2] or lexicostatistic dating, is based on the assumption that lan-

[1] A preliminary form of this paper was presented at the Summer Meeting of the Linguistic Society of America in Ann Arbor, Michigan, July 28, 1956.

[2] Among the more basic items of the rather large bibliography on Lexicostatistics are: Morris Swadesh, Lexico-Statistic Dating of Prehistoric Ethnic Contacts with Special Reference to North American Indians and Eskimos, PAPS 96.452–63 (1952); Robert B. Lees, The Basis of Glottochronology, Language 29.113–27 (1953); Morris Swadesh, Towards Greater Accuracy in Lexicostatistic Dating, IJAL 21.121–37 (1955); A. L. Kroeber, Linguistic Time Depth Results So Far and Their Meaning, IJAL 21.91–104 (1955); Harry Hoijer, Lexicostatistics: A Critique, Language 32.49–60 (1956); Sarah C. Gudschinsky, Three Disturbing Questions Concerning Lexicostatistics, IJAL 22.212–3 (1956). An excellent summary of lexicostatistic methods has been made by Sarah C. Gudschinsky in Word 12.175–210 (1956). Further bibliography may be readily found by referring to these.

guages change at a uniform and determinable rate, at least as regards certain basic lexical items, and further that this rate of change is the same for all languages. Thus, once the rate of change has been determined, one might be able to utilize a standardized test list and a mathematical formula either 1) to discover how far apart in time two samples of a given language are (to determine the age of a document, for example, by a comparison of its vocabulary with a present day or other datable sample) or 2) to determine the divergence date of two languages having a common origin. It is the purpose of this paper to examine the validity of results obtained by the second of these processes.

Morris Swadesh, utilizing a pilot study based on thirteen languages, has devised a test list[3] of one hundred items having a constant, or index, of retention of 85.4 per cent per millenium. The formula to be applied to problems involving the divergence time of two related languages is:

$$t = \log. C + 2 \log. r$$

where t is the divergence time in millenia, C is the percentage of cognate vocabulary in two related languages, and r is the index of retention, 85.4 per cent. Various processes will, of course, have contributed to the 'lexical decay' of 14.6 per cent per millenium.

This method has been applied during the past few years by a number of scholars who have seen in it an instrument useful in determining the divergence dates of languages (such as those in the American Indian groups) for which we have little or no documented history. Divergence times estimated by lexicostatistic dating run as high as 6.5 millenia.[4]

The mathematical procedures of lexicostatistic dating have been examined by Robert B. Lees, who has found that the methods used are statistically sound, assuming always that the data are accurate and the basic hypothesis valid. Professor Swadesh feels that lexicostatistic dating has reached a stage of development and refinement that has made it 'almost a precision instrument.'[5]

Two criticisms of lexicostatistics call for special mention here. Harry Hoijer[6] has recently indicated that there is frequently some difficulty in finding a simple term for each of the items in Swadesh's list. Sarah C. Gudschinsky[7] has asked: 1) Can any universally useful list be devised? 2) Are the few languages on which Swadesh and Lees have based their studies sufficient basis upon which to generalize a constant index of retention for all languages? and 3) Is a 100 word test list adequate for determining time depths as great as 6.5 millenia where there remain only 6.4 per cent of cognate vocabulary?[8]

To date no one has published any study of the results obtained from lexicostatistic dating of linguistic splits in language groups for which we have enough other information to judge whether the theory is useful and valid for this purpose. The Romance languages constitute a good testing ground for lexicostatistics, as for other linguistic hypotheses, because they contain a consider-

[3] IJAL 21.121–37. This list and the constant represent a revision of a list of over two hundred items having a retention rate of 80.48 per cent as given in PAPS 96.452–63. Only the revised list and retention rate are used in this article.

[4] Swadesh, in Problems of Long-Range Comparison in Penutian, Language 32.17–41 (1956), estimates (p. 41) a time depth of nine millenia.

[5] Swadesh, IJAL 21.137.

[6] Hoijer, Language 32.49–60.

[7] Gudschinsky, IJAL 22.212–3.

[8] A. L. Kroeber, IJAL 21.91–104, also feels that estimates based on such a small number of cognates are precarious. Although there seem to be other theoretical and methodological problems, the present paper is concerned only with the practical end of examining the validity of results obtained by lexicostatistics.

able number of languages known to have a common origin during a time which can be more accurately dated than is true of any comparable group, and we have quite a fair body of documentary evidence on various stages between the divergence date and the present day.

Likewise, six of the thirteen languages on which the original study was based were Romance. In the pilot study these were each compared with some stage of Latin, but not one with another to determine whether the divergence dates arrived at by lexicostatistics would give any date that would at all correspond to that of the Latin used by Swadesh as the proto-language, or common original of his Romance samples. In other words, the standard mathematical 'proof' consisting of running one's end product back through the original equation has never been undertaken.

One might object that the Romance languages constitute a special case, due not only to their contiguity, but likewise to a prolonged period of cultural intercommunication. If this is true, the Romance languages are indeed poor ones to have selected for what amounts to half of the original study to determine the index of retention. Furthermore, one can make such objections only because we have rather good documentation concerning the cultural history of the Romance speaking peoples, whereas lexicostatistic dating has been used without hesitation to determine the divergence dates of languages for which we have no such documentation to indicate whether the possibilities for such cultural contamination were greater or less than for the Romance group.

According to Swadesh,[9] 'The rules for filling in the list for each language may be stated as follows: a) Try to find one simple equivalent for each item by disregarding specialized and bound forms and the less common of two equivalents. b) Use a single word or element rather than a phrase, even though the meaning may be broader than that of the test item. c) Where it is impossible to find a single equivalent, omit the form.' The forms for each of the Romance languages, obtained from informants, are given below. For methodological reasons a slightly different order of the list is given from that used by Swadesh, but the items themselves are the same. The divergence dates obtained by subtracting divergence time from 1956, the year in which the data were obtained, are given in Table I.

[9] Swadesh, PAPS 96.457.

CHART I

Languages Compared	*Per cent Cognate*	*Divergence Time Depth in Millenia*	*Divergence Date (A.D.)*
Spanish-Portuguese	.89	.370	1586
Italian-French	.89	.370	1586
Italian-Catalan	.87	.440	1516
Catalan-Portuguese	.85	.516	1440
Sardinian-Italian	.85	.516	1440
Italian-Portuguese	.85	.516	1440
Spanish-Catalan	.85	.516	1440
Italian-Spanish	.82	.627	1329
Sardinian-French	.80	.706	1250
French-Catalan	.80	.706	1250
Rheto-Romance-Italian	.78	.785	1171
Rheto-Romance-French	.78	.785	1171
Sardinian-Portuguese	.78	.785	1171
Rumanian-Italian	.77	.826	1130
Rheto-Romance-Catalan	.76	.867	1089
Sardinian-Spanish	.76	.867	1089
Sardinian-Catalan	.75	.911	1045
French-Portuguese	.75	.911	1045
Rumanian-French	.75	.911	1045
French-Spanish	.75	.911	1045
Sardinian-Rumanian	.74	.953	1003
Rheto-Romance-Spanish	.74	.953	1003
Rheto-Romance-Sardinian	.74	.953	1003
Rheto-Romance-Portuguese	.74	.953	1003
Rumanian-Catalan	.73	.997	959
Rumanian-Portuguese	.72	1.041	915
Rheto-Romance-Rumanian	.72	1.041	915
Rumanian-Spanish	.71	1.082	874

ENGLISH: 1. all, 2. ashes, 3. bark, 4. bite, 5. blood, 6. bone, 7. cloud, 8. come, 9. die, 10. drink, 11. ear, 12. egg, 13. eye, 14. fire, 15. fish, 16. fly, 17. good, 18. green, 19. hand, 20. I, 21. louse, 22. man, 23. name, 24. new, 25. night, 26. nose, 27. one, 28. rain, 29. root, 30. see, 31. seed, 32. skin, 33. sleep, 34. smoke, 35. star, 36. sun, 37. thou, 38. tongue, 39. tooth, 40. two, 41. water, 42. we, 43. full, 44. round, 45. moon, 46. belly, 47. big, 48. bird, 49. black, 50. cold, 51. dog, 52. dry, 53. earth, 54. eat, 55. feather, 56. foot, 57. give, 58. hair, 59. head, 60. hear, 61. heart, 62. kill, 63. know, 64. leaf, 65. liver, 66. long, 67. many, 68. meat, 69. mountain, 70. mouth, 71. neck, 72. not, 73. red, 74. sand, 75. small, 76. stone, 77. tail, 78. tree, 79. warm, 80. white, 81. woman, 82. yellow, 83. horn, 84. knee, 85. say, 86. burn, 87. road, 88. claw, 89. person, 90. swim, 91. sit, 92. walk, 93. this, 94. that, 95. lie, 96. stand, 97. who, 98. what, 99. fat, 100. breast.

FRENCH: 1. tout, 2. cendre, 3. écorce, 4. mordre, 5. sang, 6. os, 7. nuage, 8. venir, 9. mourir, 10. boire, 11. oreille, 12. oeuf, 13. oeil, 14. feu, 15. poisson, 16. voler, 17. bon, 18. vert, 19. main, 20. je, 21. pou, 22. homme, 23. nom, 24. nouveau, 25. nuit, 26. nez, 27. un, 28. pluie, 29. racine, 30. voir, 31. semence, 32. peau, 33. dormir, 34. fumée, 35. étoile, 36. soleil, 37. tu, 38. langue, 39. dent, 40. deux, 41. eau, 42. nous, 43. plein, 44. rond, 45. lune, 46. ventre, 47. grand, 48. oiseau, 49. noir, 50. froid, 51. chien, 52. sec, 53. terre, 54. manger, 55. plume, 56. pied, 57. donner, 58. cheveu, 59. tête, 60. entendre, 61. coeur, 62. tuer, 63. savoir, 64. feuille, 65. foie, 66. long, 67. beaucoup, 68. viande, 69. montagne, 70. bouche, 71. cou, 72. ne, 73. rouge, 74. sable, 75. petit, 76. pierre, 77. queue, 78. arbre, 79. chaud, 80. blanc, 81. femme, 82. jaune, 83. corne, 84. genou, 85. dire, 86. brûler, 87. route, 88. griffe, 89. personne, 90. nager, 91. s'asseoir, 92. marcher, 93. ceci, 94. cela, 95. être étendu, 96. être debout, 97. qui, 98. que, 99. graisse, 100. poitrine.

SPANISH: 1. todo, 2. ceniza, 3. corteza, 4. morder, 5. sangre, 6. hueso, 7. nube, 8. venir, 9. morir, 10. beber, 11. oreja, 12. huevo, 13. ojo, 14. fuego, 15. pez, 16. volar, 17. bueno, 18. verde, 19. mano, 20. yo, 21. piojo, 22. hombre, 23. nombre, 24. nuevo, 25. noche, 26. nariz, 27. uno, 28. lluvia, 29. raíz, 30. ver, 31. semilla, 32. piel, 33. dormir, 34. humo, 35. estrella, 36. sol, 37. tu, 38. lengua, 39. diente, 40. dos, 41. agua, 42. nosotros, 43. lleno, 44. redondo, 45. luna, 46. barriga, 47. grande, 48. pájaro, 49. negro, 50. frío, 51. perro, 52. seco, 53. tierra, 54. comer, 55. pluma, 56. pie, 57. dar, 58. pelo, 59. cabeza, 60. oír, 61. corazón, 62. matar, 63. saber, 64. hoja, 65. hígado, 66. largo, 67. muchos, 68. carne, 69. montaña, 70. boca, 71. pescuezo, 72. no, 73. colorado, 74. arena, 75. pequeño, 76. piedra, 77. rabo, 78. árbol, 79. caliente, 80. blanco, 81. mujer, 82. amarillo, 83. cuerno, 84. rodilla, 85. decir, 86. quemarse, 87. camino, 88. garra, 89. persona, 90. nadar, 91. sentarse, 92. caminar, 93. este, 94. aquel, 95. estar recostado, 96. estar de pie, 97. quien, 98. que, 99. grasa, 100. pecho.

RUMANIAN: 1. tot, 2. cenuşă, 3. scoarţă, 4. muşca, 5. sânge, 6. os, 7. nor, 8. veni, 9. muri, 10. bea, 11. ureche, 12. ou, 13. ochiu, 14. foc, 15. peşte, 16. zbura, 17. bun, 18. verde, 19. mână, 20. eu, 21. păduche, 22. om, 23. nume, 24. nou, 25. noapte, 26. nas, 27. un, 28. ploaie, 29. rădăcină, 30. vedea, 31. sămânţă, 32. piele, 33. dormi, 34. fum, 35. stea, 36. soare, 37. tu, 38. limbă, 39. dinte, 40. doĭ, 41. apă, 42. noi, 43. plin, 44. rotund, 45. lună, 46. pântece, 47. mare, 48. pasăre, 49. negru, 50. rece, 51. câĭne, 52. uscat, 53. pământ, 54. mânca, 55. pană, 56. picior, 57. da, 58. păr, 59. cap, 60. auzi, 61. inimă, 62. omori, 63. şti, 64. foaie, 65. ficat, 66. lung, 67. mulţi, 68. carne, 69. munte, 70. gură, 71. gât, 72. nu, 73. roşie, 74. nisip, 75. mic, 76. piatră, 77. coadă, 78. copac, 79. cald, 80. alb, 81. femeie, 82. galben, 83. corn, 84. genunchiu, 85. zice, 86. arde, 87. drum, 88. ghiară, 89. om, 90. înota, 91. şedea, 92.

umbla, 93. aceasta, 94. aceasta, 95. fi culcat, 96. sta in piciore, 97. cine, 98. ce, 99. grăsime, 100. piept.

CATALAN: 1. tot, 2. cendra, 3. escorça 4. mossegar, 5. sang, 6. os, 7. núvol, 8. venir, 9. morir, 10. beure, 11. orella, 12. ou, 13. ull, 14. foc, 15. peix, 16. volar, 17. bo, 18. verd, 19. mà, 20. jo, 21. poll, 22. home, 23. nom, 24. nou, 25. nit, 26. nas, 27. un, 28. pluja, 29. arrel, 30. veure, 31. llavor, 32. pell, 33. dormir, 34. fum, 35. estel, 36. sol, 37. tu, 38. llengua, 39. dent, 40. dos, 41. aigua, 42. nosaltres, 43. ple, 44. rodó, 45. lluna, 46. ventre, 47. gran, 48. ocell, 49. negre, 50. fred, 51. gos, 52. sec, 53. terra, 54. menjar, 55. ploma, 56. peu, 57. donar, 58. cabell, 59. cap, 60. sentir, 61. cor, 62. matar, 63. saber, 64. fulla, 65. fetge, 66. llarg, 67. molts, 68. carn, 69. muntanya, 70. boca, 71. coll, 72. no, 73. vermell, 74. sorra, 75. petit, 76. pedra, 77. cua, 78. arbre, 79. calent, 80. blanc, 81. dona, 82. groc, 83. corn, 84. genoll, 85. dir, 86. cremar, 87. carretera, 88. garra, 89. home, 90. nedar, 91. seure, 92. caminar, 93. això, 94. allò, 95. jeure, 96. estar dret, 97. qui, 98. què, 99. greix, 100. pit.

ITALIAN: 1. tutto, 2. cenere, 3. corteccia, 4. mordere, 5. sangue, 6. osso, 7. nuvolo, 8. venire, 9. morire, 10. bere, 11. orecchio, 12. uovo, 13. occhio, 14. fuoco, 15. pesce, 16. volare, 17. buono, 18. verde, 19. mano, 20. io, 21. pidocchio, 22. uomo, 23. nome, 24. nuovo, 25. notte, 26. naso, 27. uno, 28. pioggia, 29. radice, 30. vedere, 31. seme, 32. pelle, 33. dormire, 34. fumo, 35. stella, 36. sole, 37. tu, 38. lingua, 39. dente, 40. due, 41. acqua, 42. noi, 43. pieno, 44. tondo, 45. luna, 46. ventre, 47. grande, 48. ucello, 49. nero, 50. freddo, 51. cane, 52. secco, 53. terra, 54. mangiare, 55. penna, 56. piede, 57. dare, 58. capelli, 59. testa, 60. sentire, 61. cuore, 62. uccidere, 63. sapere, 64. foglia, 65. fegato, 66. lungo, 67. molti, 68. carne, 69. montagna, 70. bocca, 71. collo, 72. non, 73. rosso, 74. sabbia, 75. piccolo, 76. pietra, 77. coda, 78. albero, 79. caldo, 80. bianco, 81. donna, 82. giallo, 83. corno, 84. ginocchio, 85. dire, 86. brucciare, 87. strada, 88. artiglio, 89. persona, 90. nuotare, 91. sedere, 92. camminare, 93. questo, 94. quello, 95. giacere, 96. stare in piede, 97. chi, 98. che, 99. grasso, 100. petta.

SARDINIAN: 1. tottu, 2. chisina, 3. iscorza, 4. mossigare, 5. sàmbene, 6. ossu, 7. nue, 8. bènnere, 9. mòrrere, 10. bíere, 11. urìga, 12. óu, 13. ogu, 14. fogu, 15. pische, 16. bolare, 17. bonu, 18. virde, 19. manu, 20. deo, 21. priogu, 22. òmine, 23. lúmene, 24. nóu, 25. notte, 26. nasu, 27. unu, 28. proga, 29. raighina, 30. bìere, 31. sèmene 32. pedde, 33. drumire, 34. fumu, 35. istedda, 36. sole, 37. túe, 38. limba, 39. dente, 40. duos, 41. abba, 42. nos, 43. prenu, 44. tundu, 45. luna, 46. ventre, 47. mannu, 48. puzone, 49. nièddu, 50. frittu, 51. cane, 52. siccu, 53. terra, 54. mandigare, 55. pinna, 56. pe, 57. dare, 58. pilos, 59. conca, 60. intèndere, 61. coro, 62. occhire, 63. ischire, 64. foza, 65. figadu, 66. longu, 67. meda, 68. petta, 69. muntagna, 70. bucca, 71. trugu, 72. no, 73. rugu, 74. arena, 75. picocu, 76. pedra, 77. coa, 78. àrvure, 79. cardu, 80. biancu, 81. fèmina, 82. grogu, 83. corru, 84. enugu, 85. nàrrere, 86. brusiare, 87. istrada, 88. franca, 89. pessone, 90. nadare, 91. sèzzere, 92. camminare, 93. custu, 94. cuddu, 95. est crocadu, 96. istare a pe, 97. chie, 98. chi, 99. rassu, 100. pettus.

PORTUGUESE: 1. todo, 2. cinzas, 3. cortex, 4. morder, 5. sangue, 6. osso, 7. nuvem, 8. vir, 9. morrer, 10. beber, 11. orelha, 12. ovo, 13. ôlho, 14. fogo, 15. peixe, 16. voar, 17. bom, 18. verde, 19. mão, 20. eu, 21. piolho, 22. homem, 23. nome, 24. novo, 25. noite, 26. nariz, 27. um, 28. chuva, 29. raiz, 30. ver, 31. semente, 32. pele, 33. dormir, 34. fumo, 35. estrêla, 36. sol, 37. tu, 38. língua, 39. dente, 40. dois, 41. água, 42. nos, 43. cheio, 44. redondo, 45. lua, 46. barriga, 47. grande, 48. pássaro, 49. presto, 50. frio, 51. cão, 52. sêco, 53. terra, 54. comer, 55. pena, 56. pé, 57. dar, 58. cabelo, 59. cabeça, 60. ouvir, 61. coração, 62. matar, 63. saber, 64. fôlha, 65. fígado, 66. longo, 67. muitos, 68. carne, 69. montanha, 70. bôca, 71. pescoço, 72.

não, 73. vermelho, 74. areia, 75. pequeno, 76. pedra, 77. rabo, 78. árvore, 79. quente, 80. branco, 81. mulher, 82. amarelo, 83. corno, 84. joelho, 85. dizer, 86. queimar, 87. estrada, 88. garra, 89. homem, 90. nadar, 91. sentarse, 92. caminhar, 93. êste, 94. aquele, 95. estar deitado, 96. estar em pe, 97. quem, 98. que, 99. banha, 100. peito.

RHETO-ROMANCE: 1. tut, 2. tschendra, 3. scorsa, 4. mordé, 5. saun, 6. fes, 7. nibla, 8. vanyí, 9. murí, 10. báiber, 11. urelya, 12. eyf, 13. ely, 14. fiuk, 15. pesch, 16. skulá, 17. bíen, 18. vert, 19. maun, 20. yu, 21. pluly, 22. um, 23. num, 24. neyf, 25. notg, 26. nas, 27. in, 28. pleyvya, 29. risch, 30. vazáy, 31. sem, 32. píal, 33. durmí, 34. fem, 35. steyla, 36. sulely, 37. te, 38. líunga, 39. díant, 40. dus, 41. awa, 42. nus, 43. playn, 44. radón, 45. lyina, 46. vyántar, 47. gront, 48. utschí, 49. nayr, 50. frayt, 51. tgaun, 52. sec, 53. tyara, 54. malyé, 55. plema, 56. pay, 57. da, 58. tgavely, 59. tgau, 60. udí, 61. cor, 62. matsá, 63. savây, 64. fely, 65. dir, 66. lyunk, 67. bye, 68. tgarn, 69. kuelm, 70. buka, 71. kulíets, 72. betg, 73. tgetschen, 74. sablun, 75. pints, 76. crapa, 77. cua, 78. plonta, 79. tgaut, 80. alf, 81. femna, 82. mélan, 83. tgern, 84. ganúlya, 85. di, 86. arder, 87. sulada, 88. branca, 89. karstgáun, 90. – – –, 91. sey, 92. – – –, 93. kwel, 94. tschel, 95. – – –, 96. – – –, 97. tgi, 98. tge, 99. píun, 100. pets.

It will be seen that the greatest divergence time, that found by comparing Sardinian with Rheto-Romance, is in the ninth century A.D., well after the early Romance documents which already show a considerable degree of divergence. Equally improbable are the dates for Proto-Italo-Western (Vulgar Latin) of 1003 A.D. (actually within the early literary period of several of the languages), and Proto-Ibero-Romance of 1440 A.D.[10] The figure of 75 percent for Rumanian-French is likewise amusing when it is recalled that Swadesh found a percent of 71.6 for Rumanian-Latin, and of 74.2 for French-Latin, and that he considers (IJAL 21.130–1) these two cases, along with 'Athenian,' to be the most accurate ones in his revised pilot study.

Although I agree with most Romance scholars, and other historical linguists, in hesitating to specify precise dates for linguistic splits, there is no reason to disagree with a date of around 100 A.D. as the latest time when there was any remaining phonemic uniformity, and it is likewise generally accepted that there was already some divergence in both morphology and VOCABULARY as early as 250 to 200 B.C.[11] Thus the Romance languages have been diverging lexically for close to 2.2 millenia, and the figure of 1.08 millenia obtained by the use of lexicostatistics is too far from known facts to indicate that this method of dating linguistic splits has any usefulness or validity even for the languages upon which it is based.

UNIVERSITY OF KENTUCKY

PROTOPOPOLOCA INTERNAL RELATIONSHIPS

ERIC P. HAMP

1. María Teresa Fernández de Miranda, in her Reconstrucción del protopopoloca, Revista mexicana de estudios antropológicos 12.61–93(1951), has given us an excellent ground-breaking essay at the systematic and solidly based classification of these cognate languages. Her methodology for the recovery of the greater part of the phonemic pattern of the parent language is admirable, and the results are highly trustworthy.

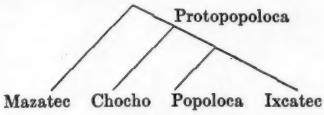
At the end of the article (93) the author

[10] The stemma used by Robert A. Hall, Jr., on page 24 of his article The Reconstruction of Proto-Romance, Language 26.6–27, gives a suitable idea of the subgroups within the Romance languages. Proto-Italo-Western is a reconstructible original of all the languages cited here except Sardinian (Southern Romance) and Rumanian (Eastern Romance). Proto-Ibero-Romance is the reconstructible original of Spanish, Portuguese and Catalan, each well attested long before the 'Age of Discovery'.

[11] Hall, Language, 26.19–20.

points out that Mazatec is more distantly related to Ixcatec, Popoloca, and Chocho than any of the latter three is to any other, because Mazatec cannot be immediately accommodated by the reconstruction that successfully accounts for the others. Miss Fernández is clearly correct.

She then goes on to say: 'Los idiomas más semejantes entre sí parecen ser el ixcateco y el popoloca, los más alejados el ixcateco y el chocho, y éste el más cercano al popoloca; pues del total de formas utilizadas para este trabajo, el ixcateco y el popoloca mostraron semejanza en 43 casos, el chocho y el popoloca en 35 y el ixcateco y el chocho en 32.' Accordingly, the stemma of relationships would appear as:



2. Now, it is not quite clear from Miss Fernández' exposition in what sense we are to understand 'semejanza', as she uses it. The only criterion for genetic proximity consists in the recognition of a decisive set, either in number or in structural placement, of shared structural innovations; and these must be innovations by addition or replacement, rather than by loss.[1] Let us inspect the phonemic correspondences set up by Miss Fernández.

3.1. Possible Ixcatec-Popoloca innovations (the numbers after each correspondence refer to Miss Fernández' cognate sets).

IP1.	Ch. d	Ix. s	Po. s	*(ð...)st 103, 106, 199, 188.
IP2.	d	ɬ²/ɬ² / t²³	s³ / t³	*ɬ(?) 95, 100, 119, 125.
IP3.	čhV	hnV	hnV	*čhnV 5.
IP4.	ñ-	ñ-/n(į)-	n-/ni-	*ñ- 3, 116, 131, 133, 190.

[1] For a more explicit statement of the principle, see E. P. Hamp, Morphological Correspondences in Cornish and Breton, The Journal of Celtic Studies 2.5–24 (1953), esp. 8–9.

[2] Before front vowel.

[3] Before back vowel.

Remarks: IP4 looks superficially promising, but it will be seen to be at least balanced out by CP11 below. What is more, on closer inspection (which any interested reader may pursue for himself) the only shred of shared innovation turns out to be the correspondence n(į)-:n-; and here we see the Popoloca phenomenon dovetails more exactly with the general loss of the palatal distinction, and the concomitant collapse of these (*tʸ, *ñ) with the dentals in Popoloca. IP1 seems to involve differential dissimilations; while we see an undeniable agreement, the change is isolated and does not structure with any larger group. IP2 looks interesting, particularly because of the matching dependence on a following vowel; but Miss Fernández remarks (67) that this correspondence is not very sure because of other problems in the same forms. We are therefore reduced, for Ixcatec-Popoloca, to IP1 and IP3.[4]

3.2. Possible Chocho-Ixcatec innovations.

CI1.	Ch. th	Ix. th	Po. hñ	*thñ	45.
CI2.	th	tʸh	hn	*tʸhn	57.
CI3.	h	h	kh	*kh	39, 123.
CI4.	š/ř/ď/r	ɬ/č	r/th	*r	12, 13, 65, 104, 143, 168, 200.
CI5.	∅	∅	-ʔ	*-ʔ	Many exx.

Remarks: CI5 is an innovation by loss, and is therefore not very significant. CI4 is so complex, and in some respects unclarified, that the area of putative agreement is scarcely important. CI1 and CI2 support one another in a neatly patterned way that is most convincing; but for the rarity of their occurrence, this is the kind of agreement that is important for genetic indications. CI3 is noteworthy, since Chocho and Ixcatec in general tend to preserve a cluster of stop plus h intact; clearly the velar cluster (was */h/*[x]?) behaved differently.

[4] An apparent agreement (but by loss) is Ch. y, Ix. ď? (before back V) č?y (before central V), Po. č?→*č?y: 14, 16, 41, 51. It should be noted that these all occur before u and a. I see no reason why the correspondence may not be equally well captioned *ď?.

In this connection, we find support in the parallel: CI3a. Ch. h Ix. šh Po. kh *škh 135,166. But we shall see below that this last is susceptible of a different interpretation.

one shared item, namely, loss of *s/š(C). The other aspects of behavior in CP1 (ie. č as against t) are not strictly parallel. CP2 and CP6, while not strong points, are mildly

CP1.	Ch. č/t/tˀ	Ix. št	Po. t/č	*-št-	46, 112, 151
CP2.	-g-	-k-	-g-	*-k-	31, 134, 141.
CP3. (=CI3a)	h	šh	kh	*škh	135, 166.
CP4.	k	sk	k	*sk	7, 164.
CP5.	k	šk	k	*šk	20, 34, 74, 89, 113, 124, 196.
CP6.	č	c	č	*(ei/u)	101.
CP7.	h	šh	hy	*šh	130, 187.
CP8.	h/Ø	-h-	h/Ø	*-h-	2, 3, 22, 77, 79, 116, 121, 188.
CP9.	hn/n(i)	hn	hn-, -n-, n(i)-	*hn	87, 89, 120, 186, 192.
CP10.	Ø	hñ	VØV	*hñ	48.
CP11.	-ñ-	-ñ-	-n-	*-n-	85
	-n(i)-	-ñ-	-ñ-		113
	-VØ-	-ñ-	VØV		61.
CP12.	-i-	-į-	-i-	*-į-	73, 79, 153.
CP13.	-i-	-a-	-i-	*-e-	11, 26.
CP14.	-u-	-a-	-u-	*-o-	14, 59, 90, 102, 178.
CP15.	e	a	i	*ɛ	17, 36 (doubtful).
CP16.	u	uu	u	*uu	122
	e	ee	e	*ee	81, 127, 129.
	a	aa	a	*aa	7, 62, 83.

Chocho-Ixcatec is thus supported by CI1, CI2, CI3, and perhaps CI3a.

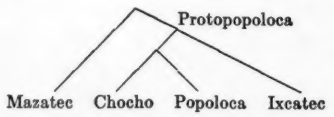
3.3. Possible Chocho-Popoloca innovations.

Remarks: The sheer number of agreements is upheld by the pervasive occurrence throughout the structure in both consonants and vowels. CP16 is an innovation by loss, and is not important. CP15 is a doubtful correspondence, but if true perhaps shows a parallel treatment (raising) of old allophones. CP13 and CP14 are highly important; they are parallel both in detail of behavior and in their symmetrical properties. CP12 is isolated. CP11 shows striking parallel detailed in allophonic development. CP10 is a loss in part, yet involves these sensitive nasal and h phenomena that are paralleled elsewhere; thus CP8, CP9, and CP10 support one another mutually. CP1, CP3, CP4, CP5, and CP7 all involve loss of initial clustered sibilants. Though this is a loss, it is significantly patterned; to recognize its unitary behavior we will regard these collectively as supporting by their similar changes in phonemic allegiance of old allophones.

We may summarize the significant shared innovations: CP2, CP6, (CP8–10), CP11, CP13, CP14, and *s/š(C).

3.4. We have found for Ixcatec-Popoloca two noteworthy agreements involving rather trivial clusters. Choco-Ixcatec yields perhaps four agreements, again involving some rather less-than-pervasive clusters. After severe pruning, Chocho-Popoloca shows seven agreements involving an important part of the vowel system, the allophones of several proto-consonants, as well as a large and homogeneous set of clusters.

4. Good methodology leads us unmistakably to associate most closely Chocho and Popoloca. The stemma therefore now appears as:



A decision on the diffusional relationships amongst these languages, arising from contiguity or long symbiosis after dialect split, must be made on the basis of a different set of procedures (including inspection of residues), and cannot be approached from the present evidence.

Considering that the above conclusion was reached on the basis of phonology alone, the outcome is gratifyingly clear-cut. It will be interesting to see what the morphology shows.

5. It remains to mention the main distinguishing features of Proto-Chocho-Popoloca. */-k-/ had an allophone *[g]; */ɸ/ had an allophone approaching *[ɸ] before */u/; */h/ and */ñ/ had certain weakened allophones; */e o/ had faulty distributions, since they did not occur internally; */s/ and */š/ no longer occurred before certain consonants. Taking CI3 into account, Protopopoloca */škh/ had already become */kh/, but this had not yet developed into */h/, as we find it in Chocho. This chain of events now demonstrates that CI3 is not really a true common innovation, but a chance later convergence. Thus the acceptable arguments for Chocho-Ixcatec are still further reduced. By elaborate argument and more data no doubt further apparent agreements could be set aside.

UNIVERSITY OF CHICAGO

THE PLACE OF ISLAND CARIB WITHIN THE ARAWAKAN FAMILY

DOUGLAS TAYLOR

If the Arawakan affiliation of Island Carib should be accepted, there then arises the question as to its place within this linguistic family or stock. A geographical answer is simple enough: Island Carib was spoken, at the time of contact, in those islands lying between the Taino area (beginning at Puerto Rico) to the north-west and the Lokono area (beginning at Trinidad) to the south and south-east. It survives to-day as the everyday home speech of some 30,000 'Caribs' (to local English-speaking populations) or 'Morenos' (to local Spanish-speaking populations) whose settlements are scattered along the shores of the Gulf of Honduras, descendants of some 5000 'Black Charaibes' who were deported from the island of St. Vincent to Roatán in 1797. A purely linguistic answer is not yet possible because all would-be classifiers of Arawakan so far have based at least their major subgroupings on geographic areas; a procedure which may prove to be linguistically justified more often than not, but which might have resulted, had it been applied to Indo-European, in classifying Rumanian dialects under Slavic and Bohemian dialects under West Germanic. In what follows I can but trace the history and suggest the nearest relationships of Island Carib in so far as these may be known or plausibly surmised from data available to me here.

The British Honduran dialects—i.e., those of Stann Creek, Hopkins, Seine Bight, Punta Gorda and Barranco—undoubtedly differ slightly from one another, and perhaps somewhat more, as a whole, from Guatemalan and Honduranian dialects such as, for example, those of Livingston, Truxillo, or the island of Roatán. But one Central American 'Carib' or 'Moreno' has no difficulty in understanding or in making himself understood by another; although either may use some words and expressions that are unfamiliar to the other, and is almost sure to betray his place of origin or upbringing by his 'accent'. These differences have most probably arisen within the past 150 years; for local varieties of Vincentian could hardly have been perpetuated as such subsequent to the deportation.

In St. Vincent and in Dominica the old language lingered on until about 1920; and there is evidence that while the dialects of these two islands differed phonologically to a considerable extent, their speakers remained mutually intelligible to the end. For Dominican we have a Catéchisme, a Dictionnaire and a Grammaire by Fr. Raymond

Breton from the middle of the XVIIth century, and an outline by Dr. Joseph Numa Rat from the end of the XIXth century; but no record of Vincentian or of any other Lesser Antillean dialect has ever been published. This is unfortunate; for it is not always possible to decide which discrepancies between the Central American and Dominican dialects go back to St. Vincent, or which of them might be imputable to the exceptionally strong Negro community in the latter island, to whom the 'Black Caribs' as a distinct group owe their origin.

Very puzzling are the numerous cases of what appears to be irregular correspondence such as, for example (DIC = Dominican Island Carib, CAIC = Central American Island Carib), DIC ísiga : CAIC íciga *to give* and DIC ísira : CAIC ígira *to leave*, but DIC sísira : CAIC sísira (*shaman's*) *rattle*. That the older Vincentian dialect did not possess the hushing sibilant affricate here written /c/ is indicated by the old loanwords, CAIC mésu *cat* (Sp. micho) and gusíu *knife* (Sp. cuchillo); and the lucky preservation of a XVIIth-century Vincentian recording, keléatibou *es-tu venu?*, suggests how it may have come into being. Compare (recent < older) DIC solohátibu < sile(h)átibu : CAIC colóhadibu < kelé(h)atibu *art thou come* or *thou art come* (a form of greeting). On the other hand, and whatever its cause (dialect mixture?) this curious and unpredictable alternation of stop, sibilant and—where it exists—affricate is common in this position (and rarely elsewhere) among other members of the Arawakan stock, as may be seen from the following examples (L = Lokono, G = Goajiro), DIC árisi : CAIC árigi : L adiki : G aciki *after*, DIC arígai : CAIC arígai : L adike : G ace?e *ear*, DIC égrigo : CAIC ógorogo : L isiroko : G i?iruku *flesh*, DIC ísiri : CAIC ígiri : L isiri : G icii *nose*, DIC asíga : CAIC acíga : L atika *to dig*, (but DIC ítiga : CAIC ídiga : L itika *excrement*).[1]

In all other respects the XVIIth-century dialect of Dominica and the present-day dialect of Central America are very close to one another (I find 94 cognates out of 100—or 179 out of 196—pairs of items belonging to basic vocabulary compared), and distinguished from their one-time neighbours, Taino and Lokono, by more than the possession of a common fund of borrowings from Cariban Kalina. So, Island Carib has, as have perhaps most Arawakan languages, a pronominal prefix of first person singular in /n/, n- *I*, *my* in place of Taino and Lokono dA-, Goajiro and Paraujano tA-; e.g., IC nitíąų : Taino datiao *my compeer*, IC núbara : L dabada : G tapatau *my (finger-)nails*, (IC /r/ regularly corresponding to L /d/ as well as to L /r/). At least the stem (iti-) of IC ítiburi *hair* has cognate equivalents in a number of Arawakan languages, but L -bara : G -wala *hair* probably have none other. IC -iuma *mouth* has cognate equivalents (such as numa, luma, nama, anuma, numaku, eomako) in at least thirty other Arawakan languages, whereas L -leroko and G -aanika, both meaning *mouth* and presumably unrelated, appear to have none. Also L hadali *sun* appears to be unique; whereas IC kási *sun*, though less common with this meaning than forms like kamoi (Arequena, Atorai, Mandauaca, Palicur, Uirina, etc.), has cognate equivalents in Piro kaši, Guana kace, Cujisenajeri tá-kaci, Canamari ghasiri, and probably in Goajiro kaih and others. G hašicci *man* is probably unique, unless it be related to G uadili *man*, which appears to have but one cognate equivalent in Palicur awaili. The latter word may, however, have been borrowed; for it is given as a synonym of hiñeiri, which, like

[1] I have assumed here that XVIIth-century Dominican Island Carib had a k/g opposition based, as certainly was that of p/b, on aspiration and its absence; but this is not certain. The consonantal phonemes of DIC appear to have been (parentheses showing doubt): p b m t s (N) n l r k (g) h; those of CAIC: p b f m t d s n l r č k g h; and those of Lokono: f (earlier p) b m t d s n l r k h. Those of Goajiro are, according to Holmer: p m w t s n l r č š y k h ʔ.

many other Arawakan words for *man*, is cognate with its IC equivalent, eiéri.

Island Carib forms show a similar phonetic development to that of their Taino, Lokono and Goajiro cognates; yet often appear to be slightly more conservative than these latter. Compare, e.g., DIC iníñe (probably /iníe/) : CAIC -íéie : L ie : G iyee : Carutana inene : Adzaneni niñe *tongue*, IC hiáru : L hiaro : G hiieer : Guinao henau : Arequena inaru *woman*, IC iráiti : L ireti : G eeci *husband*, IC áiga : L eke : G ika : Bare nika (*anika) *eat*, IC ágai : L eke *container*: IC biąburi : L bibiti < (older) biabite : G pienci : Taino yamōkobre *four* (derived from bíama : bian < biama : piama : yamoka *two*), IC baláua : L bara : G palaa : Taino baua *sea*, DIC nhA- (/nh/ was probably one phoneme in opposition with /n/) : CAIC hA- : L nA- : G nA- *they their*, IC áuli : Taino aon *dog*, IC máli : Taino mani *peanut* (borrowed from Tupian mandubi *peanut*).

Using Morris Swadesh's 100-word test-list, I find about 63% of cognates as between Lokono and Goajiro; 49%—or, if Cariban loan-words are discarded, 56%—as between either of the latter languages and older Dominican Island Carib (slightly less for CAIC); 15% as between any of these and Ariti (Paressi). These figures are based on the comparison of complete or nearly complete lists. Unfortunately, word-lists available to me for other Arawakan languages are inadequate, permitting the comparison of from twenty-five to forty pairs of words only; and sometimes a few of these had to be rejected as doubtful.[2] For example, comparison with Island Carib showed the following proportions of cognates: Guino 18/34, Catapolitani 14/31, Piapoco 17/37, Campa 12/26, Piro 12/23, Mojo 8/29, Mehinacu 10/20; and comparison with Lokono: Guinao 18/33, Catapolitani 16/32, Piapoco 16/37, Campa 11/27, Piro 13/22, Mojo 10/30, Mehinacu 11/20. Experimentation suggested that percentages based on these short lists were too high by about one ninth; thus comparing Lokono with Goajiro for the only items of the list that are available to me for Guinao I get 71% (27/38) instead of the 63% given by the complete list; and comparing Island Carib with Goajiro for the same short list I get 55% (21/38) instead of 49%. By using the other short lists, similar results were obtained; in a few cases the percentage was too high by as much as one fifth, and in no case by less than one ninth. Accordingly, percentages based on short lists have been reduced, in the following tabulation, by one ninth.

Lok	Goa	Gui	Cat	Pia	Cam	Pir	Moj	Meh	Ariti	
49	49	47	40	41	41	46	25	44	15	Island Carib (Dominican)
	63	49	44	38	36	52	29	49	15	Lokono
		53	47	41	37	62	36	44	15	Goajiro
			62	69	57	39	40	54	30	Guinao
				73	39	57	36	62	31	Catapolitani
					41	57	44	62	35	Piapoco
						67	52	37	40	Campa
							53	49	42	Piro
								53	36	Mojo
									60	Mehinacu

Island Carib, Goajiro and Piapoco belong to McQuown's major subgroup I 'Northern Arawakan'; Lokono, Guinao and Catapolitani to II 'Northern Amazon Arawakan'; Campa and Piro to III 'Pre-Andine Arawakan'; Mojo, Mehinacu and Ariti to IV 'Southern Arawakan'. The percentages tabulated above are probably all more or less inaccurate, and a few may be positively misleading; but if even a very rough approximation to actual degrees of interrelationship should be indicated by a majority of these figures, it is evident that they are not reflected at all closely by the geographic subgrouping.

Being unacquainted with any description of an Arawakan language other than Goajiro, Lokono and Island Carib, reference to grammatical criteria would serve no useful

[2] This is particularly the case when both words in a pair contain what appear to be cognate and non-cognate elements, but there is no means of telling whether the latter are members of a compound, affixes, or have come to form a part of the stem.

purpose here. But meagre as they are, both phonological and lexical data point to a rather close relationship between Island Carib on the one hand and a still tighter and more innovating group composed of Lokono, Taino, Goajiro and Paraujano on the other. This might be shown as follows:

K.1 Northern Arawakan
- a. Island Carib
- b. TA-Arawakan
  - α. Lokono
  - β. Taino
  - γ. (N.-W. Venezuela & N.-E. Columbia)
    - (1) Goajiro
    - (2) Paraujano

DOMINICA, BRITISH WEST INDIES.

CARIB, CALIBAN, CANNIBAL[1]

DOUGLAS TAYLOR

For whatever reason, the names Carib, Caribe, Caraïbe, etc. have become a part of European languages in a way that Arawak, Aruaca, Arrouague, etc. never have. Thus 'Carib'—with its derivative 'Caribbean'—is listed in the COD and 'Caraïbe' appears in both main and historical sections of the PLI, but neither 'Arawak' nor 'Arrouague' is to be found in either. Borrowed from the Arawakan language of the Taino Indians of Haiti, among whom their models both seem to have designated the contemporary inhabitants of the Lesser Antilles, Spanish caribe and caníbal have remained more or less synonymous, though their reference has been changed to *savage, cannibal*. The Island Caribs' own name for their nation, we are told by Raymond Breton, who lived among them in Dominica between 1641 and 1653, was Callíponam) probably /kalípuna/) in the Arawakan speech of the women, Callínago in the men's Cariban vocabulary. Other names which apparently sprung from the same source (most but not all of them refer to tribes that speak or spoke a Cariban language) include: karipuna, karipina, garífuna, kalihona, kalifi-(na), galibi, karibna, karifna, kali?na, kalina, karina, etc.

These differences are not very important, for no Cariban language has, so far as I know, an l/r-opposition or more than one series of stops; and while this is not true of all Arawakan languages, the latter show a good deal of instability in this respect, both as between one language and another and as between two stages of the same language. Moreover, as de Goeje has shown, comparison makes it clear that in Cariban a labial stop has often been lost by way of spirantization, as in the case of Kalina kalina (< *kalipona); and we may therefore take it that Island Carib kalípuna is a fairly conservative form. Assuming that Spanish in this case reproduces the Taino model, and comparing kaníbal with kalípuna in the light of such other pairs of words as (Taino : Island Carib) : uanabana : ualápana *sour-sop*, mani : máli *peanut*, hyen : ịhali *manioc juice* (Lokono keheli), burén : búrẽlet *cassava griddle* (Lokono budali), aon : áụli *dog*, we may conclude that—unless they were borrowed—the two words meaning *Island Carib* are related, if not cognate; (by which I mean that their final morphemes may well be different; see below).

Now according to Breton, Island Carib -na and -ri were, respectively, plural and singular suffixes employed only with stems naming or denoting places, to form derivatives designating their inhabitants; e.g.: áisina *inhabitants of Marie Galante* (from áisi *Marie Galante*), kaérabuna *inhabitants of the leeward*, kaéraburi *inhabitant of the leeward*, from kaérabu *leeward* (*side of an island*). In Lokono (True Arawak), on the other hand, -na *human plural*, -di *male singular* and -do *non-male singular* are suffixes regularly employed with one class of noun-stems, as in kalifina *Caribs or Carib nation*, kalifidi *male Carib* and kalifido *female Carib*. Island Carib /r/ is the regular correspondent of Lokono /d/, and there can

[1] It is only fair to state that I have not had the benefit of reading what P. Henríquez Ureña has written about these words (v. IJAL 23, p. 8).

be no doubt that the two sets of suffixes are cognate; but their very limited use in the former language is sufficient to account for the petrification of -na (which later became quite obsolete) in seemingly unanalysable kalípuna.

Furthermore, in Lokono, according to de Goeje, clan names are formed by suffixing -na (or, for individual members, -di or -do as explained above) to the name of a place + -io, to that of a plant + -fo, or to that of an animal or plant + -ka. So, from haiaua *incense tree* comes haiauafona *incense-tree clan*, and from barakata *armadillo* comes barakatakana *armadillo clan* (cf. arua *jaguar* and the name aruaka *Arawak*). It therefore seems plausible to suggest that Lokono kalifi- *Carib* contains (or contained) kaoi *bitter manioc* as stem, together with a variant (or residue, whose vowel may have been assimilated to preceding /i/ after petrification) of -fo as in kaiauafona. Lokono f (< p within the past hundred years) corresponds to Dominican Island Carib p, and the complete petrification of kalípuna in the latter language is to be explained by the fact that in it *manioc* was kąhį (< *kąli). This last word, like Lokono kanekidi *manioc cuttings*, correspondences with Taino (see above) and a majority of the Arawakan names for manioc point to *kaniri or kanili as a proto-form. The final morpheme of Taino kanîbal is obscure, but may well have been cognate with Island Carib -ri and Lokono -di (see above).

So that, if the hypothesis outlined above should be correct, our 'cannibal' and 'Carib' go back to Arawakan designations describing some tribe or clan as *manioc people*. Had the original name been pejorative, as those applied to foreigners often are, it would not have been widely adopted, as it has been, by the people so designated. And this in turn suggests that those arch-enemies of historical times, the Arawaks and the Caribs, were once close friends, as their traditions indeed relate.

DOMINICA, B. W. I.

iwana ~ yuana *iguana*

DOUGLAS TAYLOR

Perhaps most native speakers of English have what amounts to a prejudice against identifying the initial and second phones of 'woo' and 'ye' as belonging to the same linear phonemes; and it is hardly surprising if similar non-English sounds have often been wrongly identified as w and y; for, as Sol Saporta has recently pointed out (in IJAL 23.110), 'surely there are at least some languages where they may be considered members of /u/ and /i/.' This appears to be or to have been the case in several Arawakan languages such as Taino, Lokono (True Arawak) and Island Carib (for the latter, v. 'Phonemes of the Hopkins Dialect of Island Carib' in IJAL 21.233–241). Such wrong identification may lead, in the case of loanwords or place-names, to the emergence of two distinct forms from what was one in the source language.

Spanish (and English) iguana *iguana* comes, we are told in the COD, 'f. Carib *iwana*'. The source language was of course Taino, which we believe to have been Arawakan, but [iųána] was most probably what the first Spanish recorder of orthographic 'iguana' heard. On the other hand, we also find yuana, with the same meaning, recorded for the same language (Taino); while de Goeje gives iuwana as the name of the same animal in modern Lokono (in which stress is not distinctive). The Island Carib of Dominica (which lies midway between what were Taino and Lokono territories) seem to have called the *iguana* by a borrowed Cariban term, but what must have been its native name appears in Breton's dictionary as the stem (-yoána-) of a derived verb: ayoánaca (niábou) (*je vais*) *à la chasse aux lézards*, and probably in the island-name: Iouánacaéra *Martinique*, which contains acaéra *île, pays*. It seems clear that the model for Spanish iguana *iguana* was, phonologically speaking, neither iwana nor yuana, but iuana; though why 'yoanna' should have

prevailed in Australia, and 'iguana' in the rest of the English-speaking world, I cannot say.

Of purely local interest is the identity of the St. Vincent place-name, Owia, with the Dominica river-name, Aóya (so spelled in Breton's dictionary, and today pronounced [aúja]—in three syllables); but it is a good illustration of the same phenomenon. The meaning of this work, /auía/ or /aúia/, is *defence, guard, protection*; and stress must have wavered or shifted (with context, locality or time) so that now /u/ and now /i/ might be heard with syllabic peak, and the other high vowel as marginal.

DOMINICA, BRITISH WEST INDIES

RECENT PUBLICATIONS

General Linguistics

Aborn, Murray, and Herbert Rubinstein. Word-class distribution in sentences of fixed length. Lg 32.666–674, 1956.

Collinge, N. E. The Limitations of Historican Phonology. Archivum Linguisticum 8.111–128, 1956.

Dyen, Isadore. Language distribution and migration theory. Lg 32.611–626, 1956.

Haas, W. Concerning Glossematics. Archivum Linguisticum 8.93–110, 1956.

Stern, Theodore. Drum and Whistle "Languages": An Analysis of Speech Surrogates. AA 59.487–506, 1957.

Thomas, Lawrence L. The Linguistic Theories of N. JA. Marr. Univ. of Calif. Pubs. in Linguistics, v. 14, 1957. 176 pp.

American Indian Languages

Bright, William. The Karok Language. Univ. of Calif. Pubs. in Linguistics, v. 13, 1957. xi + 457 pp. (lithoprint).

Hymes, D. H. Some Penutian Elements and the Penutian Hypothesis. SJA 13.69–87, 1957.

Longacre, Robert E. Proto-Mixtecan. Pub. 5, Indiana Res. Center in Anthro., Folklore, and Linguistics (Part III, IJAL v. 23, no. 4), 1957. vii + 195 pp. (lithoprint).

Schneider, David M. and John M. Roberts. Zuni Kin Terms, Zuni Daily Life. Notebook No. 3, Lab. of Anthro., Univ. of Nebraska, 1956. 137 pp. (lithoprint).

Journals Received

America Indigena, v. 17, no. 1, 1957.

American Speech, v. 32, no. 2, 1957.

Archivum Linguisticum, v. 8, fasc. 2, 1956

Boletin Indigenista, v. 17, no. 3, 1957.

Japanese Journal of Ethnology, v. 21, no. 3, [1957]. [In Japanese, with summaries of each article in English.]

Kirjath Sepher (Jerusalem), v. 22, no. 2, March, 1957.

Language, v. 32, no. 4, 1956.

The Muslim World (Hartford Seminary Foundation), v. 47, no. 2, April 1957.

Oceania, v. 27, nos. 3, 4, 1957.

PMLA, v. 72, nos. 2, 3, 4, 1957.

Revista de Filología Española (Madrid), v. 39, nos. 1–4, 1955.

Word, v. 12, no. 3, 1956.